

Effective filing date
10/10/2004
11/11/2004
12/12/2004

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 15:26:43 ; Search time 52 seconds
(without alignments)
48.902 Million cell updates/sec

Title: US-09-998-350-1

Perfect score: 45

Sequence: 1 XLYENVGMV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 44 | 97.8 | 9 | 4 | AAB48919 Generic S |
| 2 | 44 | 97.8 | 9 | 4 | AAB48917 SH2 domain |
| 3 | 44 | 97.8 | 9 | 4 | AAB48922 SH2 domain |
| 4 | 44 | 97.8 | 9 | 5 | ABG68582 Peptide G |
| 5 | 44 | 97.8 | 10 | 4 | AAB48923 SH2 domain |
| 6 | 44 | 97.8 | 10 | 4 | AAB48920 SH2 domain |
| 7 | 44 | 97.8 | 10 | 4 | AAB48926 SH2 domain |
| 8 | 44 | 97.8 | 10 | 4 | AAB48921 SH2 domain |
| 9 | 44 | 97.8 | 10 | 4 | AAB48928 SH2 domain |
| 10 | 44 | 97.8 | 11 | 2 | AAB46897 SH2 domain |
| 11 | 44 | 97.8 | 11 | 2 | AAB46896 Non-phosph |
| 12 | 44 | 97.8 | 11 | 5 | ABG68419 SH2 domain |
| 13 | 44 | 97.8 | 11 | 5 | ABG68583 Peptide G |
| 14 | 44 | 97.8 | 26 | 4 | AAB48932 SH2 domain |
| 15 | 44 | 97.8 | 26 | 4 | AAB48933 SH2 domain |
| 16 | 38 | 84.4 | 9 | 4 | AAB48918 SH2 domain |
| 17 | 38 | 84.4 | 10 | 4 | AAB48924 SH2 domain |
| 18 | 38 | 84.4 | 10 | 4 | AAB48925 SH2 domain |
| 19 | 38 | 84.4 | 10 | 4 | AAB48927 SH2 domain |
| 20 | 36 | 80.0 | 11 | 2 | AAB46899 Non-phosph |
| 21 | 36 | 80.0 | 919 | 2 | AAY10382 T cell ep |
| 22 | 35 | 77.8 | 9 | 2 | AAY10382 T cell ep |
| 23 | 35 | 77.8 | 9 | 5 | ABG80064 MHC class |
| 24 | 35 | 77.8 | 9 | 7 | ADC35620 Influenza |
| 25 | 35 | 77.8 | 11 | 2 | AAB46898 Non-phosph |

| | | | | | |
|----|----|------|------|---|---------------------|
| 26 | 35 | 77.8 | 20 | 2 | AAR49328 Influenza |
| 27 | 35 | 77.8 | 20 | 2 | AAB54715 Peptide f |
| 28 | 35 | 77.8 | 244 | 2 | AAB08004 Amino aci |
| 29 | 35 | 77.8 | 244 | 2 | AAB08053 Myrotheci |
| 30 | 35 | 77.8 | 448 | 6 | ABU19327 Protein e |
| 31 | 35 | 77.8 | 562 | 2 | AAB63588 Full leng |
| 32 | 35 | 77.8 | 562 | 5 | AAB23111 Influenza |
| 33 | 35 | 77.8 | 921 | 6 | AAC23317 Rhesus mo |
| 34 | 35 | 77.8 | 931 | 6 | AAC23313 Cynomolgu |
| 35 | 34 | 75.6 | 84 | 6 | ADA08462 Human AFA |
| 36 | 34 | 75.6 | 86 | 6 | ADA08458 Chicken A |
| 37 | 34 | 75.6 | 86 | 6 | ADA08461 Avian AFA |
| 38 | 34 | 75.6 | 362 | 2 | AAY13465 Peptide S |
| 39 | 34 | 75.6 | 634 | 4 | AAB93557 Human pro |
| 40 | 34 | 75.6 | 815 | 6 | ADA08456 Chicken A |
| 41 | 34 | 75.6 | 3542 | 4 | AAB62142 P. falcip |
| 42 | 33 | 73.3 | 293 | 5 | ABG93283 C. albica |
| 43 | 33 | 73.3 | 310 | 6 | ABM68832 Photorhab |
| 44 | 32 | 71.1 | 15 | 2 | AAR58373 Partial T |
| 45 | 32 | 71.1 | 15 | 2 | AAR95531 Monoclonal |

ALIGNMENTS

RESULT 1
AAB48919
ID AAB48919 standard; peptide; 9 AA.
XX
AC AAB48919;
XX
DT 16-MAR-2001 (first entry)
XX
DE Generic SH2 domain cyclic peptide inhibitor, SEQ ID NO:3.
XX
KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytosolic; tumour; breast cancer; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1..9
FT /note= "The nitrogen atoms of the N-terminus and the C-terminal amide are joined via a bridging moiety, thereby cyclising the peptide"
FT
FT Misc-difference 1
FT /note= "Any naturally or non-naturally occurring amino acid except Glu"
FT
FT Modified-site 9
FT /note= "C-terminal amide"
XX
WO200073326-A2.
XX
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015201.
XX
PR 02-JUN-1999; 99US-0137187P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Roller PP, Long Y, Lung FT, King CR, Yang D;
XX
XX WPI; 2001-137633/14.
XX
PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.
XX
PS Disclosure; Page 5; 26pp; English.
XX
CC The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding

CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
 CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
 CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
 CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
 CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
 CC is either Aad or Glu. Optionally, there is a conservative or neutral
 CC amino acid substitution at either or both of Leu2 and Gly7, and
 CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
 CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
 CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
 CC which links the nitrogen atom of the N terminus to the nitrogen atom of
 CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
 CC turn conformation. The peptides, and compositions comprising the
 CC peptides, are useful for inhibiting the binding of the SH2 domain to a
 CC target protein. They are particularly useful for preventing cancer,
 CC especially breast cancer. The present sequence is a generic
 CC representation of a cyclic peptide of the invention
 XX
 XX Sequence 9 AA;

Query Match 97.8%; Score 44; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
 DB 2 LYENVGMV 9

RESULT 2
 AAB48917
 ID AAB48917 standard; peptide; 9 AA.

AC AAB48917;
 DT 16-MAR-2001 (first entry)

DE SH2 domain cyclic peptide inhibitor, SEQ ID NO:1.
 KW SH2 domain binding inhibitor; non-phosphorylated; redox stable;
 KW cytosstatic; tumour; breast cancer; cyclic.
 OS Synthetic.

Key Location/Qualifiers
 FT Modified-site 1..9
 FT /note= "The nitrogen atoms of the N-terminus and the C-
 FT terminal amide are joined via a bridging moiety, thereby
 FT cyclising the peptide"
 FT Modified-site 1
 FT /note= "Gamma-carboxyglutamic acid"
 FT Modified-site 9
 FT /note= "C-terminal amide"

XX WO200073326-A2.
 PN
 XX
 XX 07-DEC-2000.

XX 02-JUN-2000; 2000WO-US015201.
 XX
 XX 02-JUN-1999; 99US-0137187P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Rolier PP, Long Y, Lung FT, King CR, Yang D;
 XX WPI; 2001-137633/14.

XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
 XX homology 2 domain binding to target protein, useful for preventing
 XX cancer, especially breast cancer.

XX Claim 1; Page 21; 26pp; English.
 PS
 XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
 CC which bind to Src homology 2 (SH2) domains, preventing them from binding
 CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
 CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
 CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
 CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
 CC aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
 CC is either Aad or Glu. Optionally, there is a conservative or neutral
 CC amino acid substitution at either or both of Leu2 and Gly7, and
 CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
 CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
 CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
 CC which links the nitrogen atom of the N terminus to the nitrogen atom of
 CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
 CC turn conformation. The peptides, and compositions comprising the
 CC peptides, are useful for inhibiting the binding of the SH2 domain to a
 CC target protein. They are particularly useful for preventing cancer,
 CC especially breast cancer. The present sequence represents a cyclic
 CC peptide of the invention
 XX
 XX Sequence 9 AA;

Query Match 97.8%; Score 44; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
 DB 2 LYENVGMV 9

RESULT 3
 AAB48922
 ID AAB48922 standard; peptide; 9 AA.

AC AAB48922;

DT 16-MAR-2001 (first entry)

DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:7.

XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
 KW cytosstatic; tumour; breast cancer; linear precursor.

OS Synthetic.

Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Gamma-carboxyglutamic acid; the nitrogen atom of
 FT the N-terminus is joined to a ClCH2C(O) moiety"
 FT Modified-site 9
 FT /note= "The carbon atom of the C-terminus is joined to a
 FT C(CH2SH)C(O)NH2 moiety"

XX WO200073326-A2.

XX 07-DEC-2000.

XX 02-JUN-2000; 2000WO-US015201.

XX 02-JUN-1999; 99US-0137187P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Rolier PP, Long Y, Lung FT, King CR, Yang D;

XX WPI; 2001-137633/14.

XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
 XX homology 2 domain binding to target protein, useful for preventing
 XX cancer, especially breast cancer.

PT Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
PT homology 2 domain binding to target protein, useful for preventing
PT cancer, especially breast cancer.

XX Example 1; Page 13; 26pp; English.

XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
CC which bind to Src homology 2 (SH2) domains, preventing them from binding
CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn4-Val5-Gly7-Met8-
CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
CC aminoadipic acid (Aad), referred to as Adi in the specification); and Xaa3
CC is either Aad or Glu. Optionally, there is a conservative or neutral
CC amino acid substitution at either or both of Leu2 and Gly7, and
CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
CC which links the nitrogen atom of the N terminus to the nitrogen atom of
CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
CC of less than 4.0 micromolar when the target protein is Grb2 (growth
CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
CC turn conformation. The peptides' and compositions comprising the
CC peptides, are useful for inhibiting the binding of the SH2 domain to a
CC target protein. They are particularly useful for preventing cancer,
CC especially breast cancer. The present sequence represents a linear
CC precursor of a peptide of the invention

XX Sequence 9 AA;

Query Match 97.8%; Score 44; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LYENVGMV 9
DB 2 LYENVGMV 9

RESULT 4

ID ABG68582 standard; peptide; 9 AA.

XX ABG68582;

DT 07-OCT-2002 (first entry)

DE Peptide GITE #1.

XX Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
KW cytosolic; cancer; phase display; tumour; metastasis; breast cancer;
KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
KW breast disorder; oesophageal disorder; pancreatic disorder; GI;
KW prostate disorder; small intestine disorder; placental disorder;
KW colon disorder; ovary disorder; testicular disorder; lung disorder.

XX Synthetic.

XX WO200236142-A2.

XX 10-MAY-2002.

XX 05-NOV-2001; 2001WO-US047400.

XX 03-NOV-2000; 2000US-0245755P.

XX (UYVE-V) UNIV VERMONT & STATE AGRIC COLLEGE.

XX Krag DN, Pero SC, Oligino L;

XX WPI; 2002-547451/58.

PT Treatment or prophylaxis of a subject having a disorder characterized by

PT abnormal interaction of Grb7 and a Grb7 ligand, involves administering to
PT a non-phosphorylated peptide to a subject in need of the treatment.

PS Disclosure; Fig 9B; 186pp; English.

XX The invention relates to treatment or prophylaxis (M1) of a subject
CC having a disorder characterised by abnormal interaction of Grb7 (Growth
CC factor receptor-bound protein 7 and a Grb7 ligand, comprising
CC administering to a subject in need of the treatment, a non-phosphorylated
CC peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-
CC Asn) or its functional equivalent, in an amount effective to inhibit the
CC disorder. Also included are peptide antagonists/inhibitors of Grb7,
CC nucleic acids encoding the antagonists, an expression vector comprising
CC the nucleic acid, a host cell transformed or transfected with the vector,
CC screening (M2) a molecular library to identify a compound that inhibits
CC interaction between Grb7 and a peptide antagonist and a phase display
CC library comprising Grb7 antagonists. M1 is useful for prophylaxis or
CC treatment of a subject having a disorder characterised by abnormal
CC interaction of Grb7 and a Grb7 ligand, including breast or oesophageal
CC cancer, primary tumour or metastasis, or disorders in kidney, liver,
CC gonads, breast, oesophagus, pancreas, prostate, small intestine,
CC placenta, colon, ovary, testes and lung. The present sequence is a GI
CC peptide (not defined) or derivative which is used to illustrate the
CC possible structures of cyclic Grb7 antagonists

XX Sequence 9 AA;

Query Match 97.8%; Score 44; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LYENVGMV 9
DB 2 LYENVGMV 9

RESULT 5

ID AAB48923 standard; peptide; 10 AA.

XX AAB48923;

XX 16-MAR-2001 (first entry)

DE SH2 domain cyclic peptide inhibitor, SEQ ID NO:8.

XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
KW cytosolic; tumour; breast cancer; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1...10 The nitrogen atoms of the N-terminus and the C-
FT terminal amide are joined via a bridging moiety, thereby
FT cyclising the peptide

FT Modified-site 1 /label= Aad

FT Modified-site 10 /note= "C-terminal amide"

XX WO200073326-A2.

XX 07-DEC-2000.

XX 02-JUN-2000; 2000WO-US015201.

XX 02-JUN-1999; 99US-0137187P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Roller PP, Long Y, Lung FT, King CR, Yang D;


```

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Roller PP, Long Y, Lung FT, King CR, Yang D;
XX WPI; 2001-137633/14.
XX
XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
XX homology 2 domain binding to target protein, useful for preventing
XX cancer, especially breast cancer.
XX
XX Example 4; Page 14; 26pp; English.
XX
XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
XX which bind to Src homology 2 (SH2) domains, preventing them from binding
XX to phosphotyrosine (pTyr)-containing regions of target proteins. The
XX cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
XX -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
XX Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
XX aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
XX is either Aad or Glu. Optionally, there is a conservative or neutral
XX amino acid substitution at either or both of Leu2 and Gly7, and
XX optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
XX The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
XX -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
XX which links the nitrogen atom of the N terminus to the nitrogen atom of
XX the C-terminal amide. The peptides are characterised by an in vivo IC-50
XX of less than 4.0 micromolar when the target protein is Grb2 (growth
XX factor receptor-bound protein 2). On binding Grb2, the peptides have a
XX turn conformation. The peptides, and compositions comprising the
XX peptides, are useful for inhibiting the binding of the SH2 domain to a
XX target protein. They are particularly useful for preventing cancer,
XX especially breast cancer. The present sequence represents a linear
XX precursor of a peptide of the invention
XX
XX Sequence 10 AA;
XX
XX Query Match 97.8%; Score 44; DB 4; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 0.014;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 LYENVGMVY 9
XX Db |||||
XX 2 LYENVGMVY 9
XX
XX RESULT 8
XX AAB48921
XX ID AAB48921 standard; peptide; 10 AA.
XX AC AAB48921;
XX DT 16-MAR-2001 (first entry)
XX DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:5.
XX
XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
XX cytosstatic; tumour; breast cancer; linear precursor.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1
XX FT /note= "Gamma-carboxyglutamic acid"
XX FT
XX WO2000073326-A2.
XX
XX 07-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US015201.
XX
XX 02-JUN-1999; 99US-0137187P.
XX
XX

```

```

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Roller PP, Long Y, Lung FT, King CR, Yang D;
XX WPI; 2001-137633/14.
XX
XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
XX homology 2 domain binding to target protein, useful for preventing
XX cancer, especially breast cancer.
XX
XX Example 1; Page 12; 26pp; English.
XX
XX The invention relates to redox-stable, non-phosphorylated cyclic peptides
XX which bind to Src homology 2 (SH2) domains, preventing them from binding
XX to phosphotyrosine (pTyr)-containing regions of target proteins. The
XX cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
XX -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
XX Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
XX aminoadipic acid (Aad, referred to as Adi in the specification); and Xaa3
XX is either Aad or Glu. Optionally, there is a conservative or neutral
XX amino acid substitution at either or both of Leu2 and Gly7, and
XX optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
XX The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
XX -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
XX which links the nitrogen atom of the N terminus to the nitrogen atom of
XX the C-terminal amide. The peptides are characterised by an in vivo IC-50
XX of less than 4.0 micromolar when the target protein is Grb2 (growth
XX factor receptor-bound protein 2). On binding Grb2, the peptides have a
XX turn conformation. The peptides, and compositions comprising the
XX peptides, are useful for inhibiting the binding of the SH2 domain to a
XX target protein. They are particularly useful for preventing cancer,
XX especially breast cancer. The present sequence represents a linear
XX precursor of a peptide of the invention
XX
XX Sequence 10 AA;
XX
XX Query Match 97.8%; Score 44; DB 4; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 0.014;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 LYENVGMVY 9
XX Db |||||
XX 2 LYENVGMVY 9
XX
XX RESULT 9
XX AAB48928
XX ID AAB48928 standard; peptide; 10 AA.
XX AC AAB48928;
XX DT 16-MAR-2001 (first entry)
XX DE SH2 domain peptide inhibitor linear precursor, SEQ ID NO:14.
XX
XX SH2 domain binding inhibitor; non-phosphorylated; redox stable;
XX cytosstatic; tumour; breast cancer; linear precursor.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 10
XX FT /label= Aad
XX FT /note= "C-terminal amide, joined to a solid matrix"
XX FT
XX WO2000073326-A2.
XX
XX 07-DEC-2000.
XX
XX 02-JUN-2000; 2000WO-US015201.
XX
XX 02-JUN-1999; 99US-0137187P.
XX
XX

```

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Rollier PP, Long Y, Lung FT, King CR, Yang D;
 XX
 DR WPI; 1998-110340/10.
 XX
 XX Redox-stable, non-phosphorylated cyclic peptide inhibitors of the Src
 PT. homology 2 domain binding to target protein, useful for preventing
 PT cancer, especially breast cancer.
 XX
 PS Example 5; Page 15; 26pp; English.
 XX
 CC- The invention relates to redox-stable, non-phosphorylated cyclic peptides
 CC which bind to src homology 2 (SH2) domains, preventing them from binding
 CC to phosphotyrosine (pTyr)-containing regions of target proteins. The
 CC cyclic peptides are of one of the following formulae: Xaa1-Leu2-Tyr3-Glu4
 CC -Asn5-Val6-Gly7-Met8-Tyr9-NH or Xaa2-Leu2-Tyr3-Xaa3-Asn5-Val6-Gly7-Met8-
 CC Tyr9-NH where: Xaa1 is gamma-carboxy-L-glutamic acid (Gla); Xaa2 is 2-
 CC aminocadipic acid (Aad), referred to as Aad in the specification); and Xaa3
 CC is either Aad or Glu. Optionally, there is a conservative or neutral
 CC amino acid substitution at either or both of Leu2 and Gly7, and
 CC optionally one or more of Tyr3, Glu4, Val6, Met8 and Tyr9 is modified.
 CC The peptides are cyclised via a bridging moiety of the formula C(O)-CH2-Z
 CC -CH2-CHC(O)NH2, where Z is sulphur, sulphoxide, oxygen or methylene,
 CC which links the nitrogen atom of the N terminus to the nitrogen atom of
 CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
 CC turn conformation. The peptides, and compositions comprising the
 CC peptides, are useful for inhibiting the binding of the SH2 domain to a
 CC target protein. They are particularly useful for preventing cancer,
 CC especially breast cancer. The present sequence represents a linear
 CC precursor of a peptide of the invention
 XX
 SQ Sequence 10 AA;
 Query Match 97.8%; Score 44; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LYENVGMY 9
 DB. |||||
 2 LYENVGMY 9
 RESULT 10
 AAM46897
 ID AAM46897 standard; peptide; 11 AA.
 XX
 AC AAM46897;
 XX
 DT 19-JUN-1998 (first entry)
 XX
 DE G1C-S peptide.
 XX
 KW SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
 KW signal transduction protein; non-phosphorylated; inhibition; treatment;
 KW hyper-proliferative disease; human cancer.
 XX
 OS Unidentified.
 XX
 PN WO9802176-A1.
 XX
 XX 22-JAN-1998.
 XX
 XX 16-JUL-1997; 97WO-US012501.
 XX
 XX 16-JUL-1996; 96US-0021858P.
 XX
 XX (GEOU) UNIV GEORGETOWN.
 PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
 XX
 XX King CR, Sastry L, Krag D, Oligino L;
 XX
 PI WPI; 1998-110340/10.
 XX
 XX Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
 PT transducing protein - at least as well as natural phosphorylated target,
 PT particularly from treatment of cancer.
 XX
 PS Disclosure; Page 18; 39pp; English.
 XX
 CC The present sequence represents a peptide designated G1C-S. This peptide
 CC is essentially the same as a non-phosphorylated peptide, G1, that is
 CC capable of binding to the src homology 2 (SH2) domain of Grb2, except
 CC that the terminal Cys residues of G1 are replaced with Ser residues. Grb2
 CC is a signal transduction protein. The binding affinity of the present
 CC peptide with Grb2 was tested, and it was demonstrated that the disulphide
 CC bond of G1 may be important. The G1 peptide binds to the SH2 domain of
 CC Grb2 with affinity similar to, or greater than, that of a SHC
 CC phosphopeptide (AAM46895). The G1 peptide contains a tyrosine residue
 CC that has not been modified by phosphate or similar charged group. The G1
 CC peptide is used to inhibit a signal transduction process that involves
 CC binding of a phosphorylated protein or peptide to the SH2 domain of a
 CC signal transduction protein, particularly Grb2. It is used specifically
 CC for treatment of hyper-proliferative diseases, especially human cancer
 XX
 SQ Sequence 11 AA;
 Query Match 97.8%; Score 44; DB 2; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LYENVGMY 9
 DB. |||||
 3 LYENVGMY 10
 RESULT 11
 AAM46896
 ID AAM46896 standard; peptide; 11 AA.
 XX
 AC AAM46896;
 XX
 DT 19-JUN-1998 (first entry)
 XX
 DE Non-phosphorylated peptide which binds to the SH2 domain of Grb2.
 XX
 KW SHC phosphopeptide; binding; src homology 2 domain; SH2 domain; Grb2;
 KW signal transduction protein; non-phosphorylated; inhibition; treatment;
 KW hyper-proliferative disease; human cancer; cyclic.
 XX
 OS Unidentified.
 XX
 XX Key Location/Qualifiers
 FT Disulfide-bond 1..11
 XX
 PN WO9802176-A1.
 XX
 XX 22-JAN-1998.
 XX
 XX 16-JUL-1997; 97WO-US012501.
 XX
 XX 16-JUL-1996; 96US-0021858P.
 XX
 XX (GEOU) UNIV GEORGETOWN.
 PA (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
 XX
 XX King CR, Sastry L, Krag D, Oligino L;
 XX
 XX WPI; 1998-110340/10.
 XX
 XX Non-phosphorylated peptide(s) that bind Src Homology 2 domain of signal
 PT transducing protein - at least as well as natural phosphorylated target,
 PT particularly from treatment of cancer.
 XX

PS Claim 9; Page 17; 39pp; English.

XX The present sequence represents non-phosphorylated peptide, G1, that is capable of binding to the src homology 2 (SH2) domain of Grb2. Grb2 is a signal transduction protein. The G1 peptide binds to the SH2 domain of Grb2 with affinity similar to, or greater than, that of a SHC phosphopeptide (AAW4895). The G1 peptide contains a tyrosine residue that has not been modified by phosphate or similar charged group. The G1 peptide is used to inhibit a signal transduction process that involves binding of a phosphorylated protein or peptide to the SH2 domain of a signal transduction protein, particularly Grb2. It is used specifically for treatment of hyper-proliferative diseases, especially human cancer

XX

SQ Sequence 11 AA;

Query Match 97.8%; Score 44; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LYENVGMY 9
| | | | |
Db 3 LYENVGMY 10
| | | | |

RESULT 12
ABG68419
ID ABG68419 standard; peptide; 11 AA.
XX
AC ABG68419;
DT 07-OCT-2002 (first entry)
DE G1 peptide.
XX Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
KW cytosolic; cancer; phage display; tumour; metastasis; breast cancer;
KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
KW breast disorder; oesophageal disorder; pancreatic disorder; GI;
KW prostate disorder; small intestine disorder; placental disorder;
KW colon disorder; ovary disorder; testicular disorder; lung disorder.
XX
OS Synthetic.
XX
XX WO200236142-A2.
XX
PD 10-MAY-2002.
XX
PF 05-NOV-2001; 2001WO-US047400.
XX
PR 03-NOV-2000; 2000US-0245755P.
XX
XX (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PA Krag DN, Pero SC, Oligino L;
PI WPI; 2002-547451/58.
XX
XX Treatment or prophylaxis of a subject having a disorder characterized by abnormal interaction of Grb7 and a Grb7 ligand, involves administering to a non-phosphorylated peptide to a subject in need of the treatment.
XX
PS Disclosure; Page 102; 186pp; English.

XX The invention relates to treatment or prophylaxis (M1) of a subject having a disorder characterised by abnormal interaction of Grb7 (Growth factor receptor-bound protein 7 and a Grb7 ligand, comprising administering to a subject in need of the treatment, a non-phosphorylated peptide comprising a sequence (S1, Tyr-Ala-Asn and Tyr-Asp-Asn) or its functional equivalent, in an amount effective to inhibit the disorder. Also included are peptide antagonists/inhibitors of Grb7, nucleic acids encoding the antagonists, an expression vector comprising the nucleic acid, a host cell transformed or transfected with the vector, screening (M2) a molecular library to identify a compound that inhibits interaction between Grb7 and a peptide antagonist and a phage display library comprising Grb7 antagonists. M1 is useful for prophylaxis or treatment of a subject having a disorder characterised by abnormal

CC interaction between Grb7 and a peptide antagonist and a phage display library comprising Grb7 antagonists. M1 is useful for prophylaxis or treatment of a subject having a disorder characterised by abnormal interaction of Grb7 and a Grb7 ligand, including breast or oesophageal cancer, primary tumour or metastasis, or disorders in kidney, liver, gonads, breast, oesophagus, pancreas, prostate, small intestine, placenta, colon, ovary, testes and lung. The present sequence is a G1 peptide (not defined) or derivative which is used to illustrate the possible structures of cyclic Grb7 antagonists

XX

SQ Sequence 11 AA;

Query Match 97.8%; Score 44; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LYENVGMY 9
| | | | |
Db 3 LYENVGMY 10
| | | | |

RESULT 13
ABG68583
ID ABG68583 standard; peptide; 11 AA.
XX
AC ABG68583;
DT 07-OCT-2002 (first entry)
DE Peptide GLTE #2.
XX
XX Growth factor receptor-bound protein 7; Grb7; ligand; antagonist;
KW cytosolic; cancer; phage display; tumour; metastasis; breast cancer;
KW oesophageal cancer; kidney disorder; liver disorder; gonad disorder;
KW breast disorder; oesophageal disorder; pancreatic disorder; GI;
KW prostate disorder; small intestine disorder; placental disorder;
KW colon disorder; ovary disorder; testicular disorder; lung disorder.
XX
OS Synthetic.
XX
XX WO200236142-A2.
XX
PD 10-MAY-2002.
XX
PF 05-NOV-2001; 2001WO-US047400.
XX
PR 03-NOV-2000; 2000US-0245755P.
XX
XX (UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.
XX
PA Krag DN, Pero SC, Oligino L;
PI WPI; 2002-547451/58.
XX
XX Treatment or prophylaxis of a subject having a disorder characterized by abnormal interaction of Grb7 and a Grb7 ligand, involves administering to a non-phosphorylated peptide to a subject in need of the treatment.
XX
PS Disclosure; Fig 9C; 186pp; English.

XX The invention relates to treatment or prophylaxis (M1) of a subject having a disorder characterised by abnormal interaction of Grb7 (Growth factor receptor-bound protein 7 and a Grb7 ligand, comprising administering to a subject in need of the treatment, a non-phosphorylated peptide comprising a sequence (S1, Tyr-Ala-Asn, Tyr-Glu-Asn and Tyr-Asp-Asn) or its functional equivalent, in an amount effective to inhibit the disorder. Also included are peptide antagonists/inhibitors of Grb7, nucleic acids encoding the antagonists, an expression vector comprising the nucleic acid, a host cell transformed or transfected with the vector, screening (M2) a molecular library to identify a compound that inhibits interaction between Grb7 and a peptide antagonist and a phage display library comprising Grb7 antagonists. M1 is useful for prophylaxis or treatment of a subject having a disorder characterised by abnormal

of less than 4.0 micromolar when the target protein is Grb2 (grown factor receptor-bound protein 2). On binding Grb2, the peptides have a

CC the C-terminal amide. The peptides are characterised by an in vivo IC-50
 CC of less than 4.0 micromolar when the target protein is Grb2 (growth
 CC factor receptor-bound protein 2). On binding Grb2, the peptides have a
 CC turn conformation. The peptides, and compositions comprising the
 CC peptides, are useful for inhibiting the binding of the SH2 domain to a
 CC target protein. They are particularly useful for preventing cancer,
 CC especially breast cancer. The present sequence represents a cyclic
 CC peptide of the invention
 XX
 SQ Sequence 26 AA;

Query Match 97.8%; Score 44; DB 4; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.042;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LYENVGMY 9
 |||||
 Db 2 LYENVGMY 9

Search completed: July 20, 2004, 15:43:30
 Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 15:43:34 ; Search time 42 Seconds
(without alignments)
66,977 Million cell updates/sec

Title: US-09-998-350-1

Perfect score: 45

Sequence: 1 XLYENVGM Y 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1285356 seqs, 312550742 residues

Total number of hits satisfying chosen parameters: 1285356

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|------------------|
| 1 | 44 | 97.8 | 9 | 10 | US-09-998-350-1 |
| 2 | 44 | 97.8 | 9 | 10 | US-09-998-350-3 |
| 3 | 44 | 97.8 | 9 | 10 | US-09-998-350-7 |
| 4 | 44 | 97.8 | 10 | 10 | US-09-998-350-4 |
| 5 | 44 | 97.8 | 10 | 10 | US-09-998-350-5 |
| 6 | 44 | 97.8 | 10 | 10 | US-09-998-350-6 |
| 7 | 44 | 97.8 | 10 | 10 | US-09-998-350-8 |
| 8 | 44 | 97.8 | 10 | 10 | US-09-998-350-11 |
| 9 | 44 | 97.8 | 10 | 10 | US-09-998-350-14 |
| 10 | 44 | 97.8 | 11 | 14 | US-10-013-815-32 |
| 11 | 44 | 97.8 | 26 | 10 | US-09-998-350-18 |
| 12 | 44 | 97.8 | 26 | 10 | US-09-998-350-19 |
| 13 | 38 | 84.4 | 9 | 10 | US-09-998-350-2 |
| 14 | 38 | 84.4 | 10 | 10 | US-09-998-350-9 |
| 15 | 38 | 84.4 | 10 | 10 | US-09-998-350-10 |

38 84.4 10 10 US-09-998-350-12
38 84.4 10 10 US-09-998-350-13
35 77.8 9 12 US-10-367-580-48
18 35 77.8 9 12 US-10-367-593-48
19 35 77.8 9 12 US-10-367-594-48
20 35 77.8 9 12 US-10-367-654-48
21 35 77.8 9 12 US-10-367-654-48
22 35 77.8 9 12 US-10-367-658-48
23 35 77.8 9 12 US-10-367-668-48
24 35 77.8 9 16 US-10-367-674-48
25 35 77.8 9 16 US-10-777-053-366
26 35 77.8 9 16 US-10-777-053-943
27 35 77.8 9 16 US-10-777-053-958
28 35 77.8 244 15 US-10-392-301-33
29 35 77.8 448 12 US-10-292-122A-47251
30 34 75.6 79 14 US-10-246-354-7
31 34 75.6 84 14 US-10-246-354-10
32 34 75.6 86 14 US-10-424-599-170035
33 34 75.6 168 12 US-10-424-599-170035
34 34 75.6 815 14 US-10-246-354-3
35 34 75.6 1896 16 US-10-408-765A-747
36 34 75.6 3542 12 US-10-087-013-2
37 33 73.3 1234 15 US-10-369-493-13287
38 32 71.1 78 12 US-10-424-599-219681
39 32 71.1 134 16 US-10-437-963-168439
40 32 71.1 149 15 US-10-369-493-22412
41 32 71.1 162 12 US-10-424-599-205104
42 32 71.1 306 15 US-10-369-493-1088
43 32 71.1 416 12 US-10-424-599-198277
44 32 71.1 434 9 US-09-815-242-4987
45 32 71.1 434 15 US-10-369-493-5028

ALIGNMENTS

RESULT 1

US-09-998-350-1

; Sequence 1, Application US/09998350

; Publication No. US20030078368A1

; GENERAL INFORMATION:

; APPLICANT: Roller, Peter P

; APPLICANT: Long, Ya-Qiu

; APPLICANT: Lung, Feng-Di T

; APPLICANT: King, Richter C

; APPLICANT: Yang, Dajun

; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2

; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND

; TITLE OF INVENTION: SYNTHESIS AND USE

; FILE REFERENCE: 214683

; CURRENT APPLICATION NUMBER: US/09/998,350

; CURRENT FILING DATE: 2002-12-09

; PRIOR APPLICATION NUMBER: PCT/US00/15201

; PRIOR FILING DATE: 2000-06-02

; PRIOR APPLICATION NUMBER: 60/137,187

; PRIOR FILING DATE: 1999-06-02

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(1)

; OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (9)..(9)

; OTHER INFORMATION: Tyr at position 9 is an amide, i.e. C(O)NH

; FEATURE:

; NAME/KEY: misc feature

LOCATION: (1)..(9)
 OTHER INFORMATION: Xaa (Gla) and Tyr at position 9 are bridged together, making this peptide cyclic
 OTHER INFORMATION: peptide cyclic
 US-09-998-350-1

Query Match 97.8%; Score 44; DB 10; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
 |||||

Db 2 LYENVGMY 9
 |||||

RESULT 2

US-09-998-350-3
 Sequence 3, Application US/09998350
 Publication No. US20030078368A1
 GENERAL INFORMATION:

APPLICANT: Roller, Peter P
 APPLICANT: Long, Ya-Qiu
 APPLICANT: Lung, Feng-Di T
 APPLICANT: King, Richter C
 APPLICANT: Yang, Dajun

TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
 TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
 FILE REFERENCE: 214683

CURRENT APPLICATION NUMBER: US/09/998,350
 CURRENT FILING DATE: 2002-12-09
 PRIOR APPLICATION NUMBER: PCT/US00/15201
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: 60/137,187
 PRIOR FILING DATE: 1999-06-02
 NUMBER OF SEQ ID NOS: 19
 SOFTWARE: PatentIn version 3.1

SEQ ID NO 3
 LENGTH: 9

TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthetic
 FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)..(1)

OTHER INFORMATION: Xaa is any amino acid other than Glu

FEATURE:

NAME/KEY: misc.feature

LOCATION: (9)..(9)

OTHER INFORMATION: Tyr at position 9 is an amide, i.e., C(O)NH

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)..(9)

OTHER INFORMATION: Xaa and Tyr at position 9 are bridged together, making this peptide cyclic
 OTHER INFORMATION: de cyclic

US-09-998-350-3

Query Match 97.8%; Score 44; DB 10; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
 |||||

Db 2 LYENVGMY 9
 |||||

RESULT 3

US-09-998-350-7

Sequence 7, Application US/09998350
 Publication No. US20030078368A1
 GENERAL INFORMATION:

APPLICANT: Roller, Peter P
 APPLICANT: Long, Ya-Qiu

APPLICANT: Lung, Feng-Di T
 APPLICANT: King, Richter C
 APPLICANT: Yang, Dajun
 TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
 TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
 FILE REFERENCE: 214683

CURRENT APPLICATION NUMBER: US/09/998,350
 CURRENT FILING DATE: 2002-12-09
 PRIOR APPLICATION NUMBER: PCT/US00/15201
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: 60/137,187
 PRIOR FILING DATE: 1999-06-02
 NUMBER OF SEQ ID NOS: 19
 SOFTWARE: PatentIn version 3.1

SEQ ID NO 7

LENGTH: 9

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)..(1)

OTHER INFORMATION: Xaa = Gla, which is gamma-carboxy-L-glutamic acid

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)..(1)

OTHER INFORMATION: Xaa has a ClCH2C(O)- group attached

FEATURE:

NAME/KEY: misc.feature

LOCATION: (9)..(9)

OTHER INFORMATION: Tyr at position 9 has a -C(CH2SH)C(O)NH2 group attached

US-09-998-350-7

Query Match 97.8%; Score 44; DB 10; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.2e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
 |||||

Db 2 LYENVGMY 9
 |||||

RESULT 4

US-09-998-350-4

Sequence 4, Application US/09998350
 Publication No. US20030078368A1
 GENERAL INFORMATION:

APPLICANT: Roller, Peter P

APPLICANT: Long, Ya-Qiu

APPLICANT: Lung, Feng-Di T

APPLICANT: King, Richter C

APPLICANT: Yang, Dajun

TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
 TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
 FILE REFERENCE: 214683

CURRENT APPLICATION NUMBER: US/09/998,350

CURRENT FILING DATE: 2002-12-09

PRIOR APPLICATION NUMBER: PCT/US00/15201

PRIOR FILING DATE: 2000-06-02

PRIOR APPLICATION NUMBER: 60/137,187

PRIOR FILING DATE: 1999-06-02

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn version 3.1

SEQ ID NO 4

LENGTH: 10

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic

FEATURE:

```

; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys at position 10 is an amide, i.e., C(O)NH
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(10)
; OTHER INFORMATION: Xaa (Glu) and Cys are bridged together, making this peptide cyclic
; OTHER INFORMATION: C
US-09-998-350-4

```

```

Query Match          97.8%; Score 44; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Qy      2 LYENVGMVY 9
        |||||

```

```
Db      2 LYENVGMVY 9

```

RESULT 5

```

US-09-998-350-5
; Sequence 5, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid
US-09-998-350-5

```

```

Query Match          97.8%; Score 44; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Qy      2 LYENVGMVY 9
        |||||

```

```
Db      2 LYENVGMVY 9

```

RESULT 6

```

US-09-998-350-6
; Sequence 6, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid
US-09-998-350-5

```

```

Query Match          97.8%; Score 44; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Qy      2 LYENVGMVY 9
        |||||

```

```
Db      2 LYENVGMVY 9

```

RESULT 7

```

US-09-998-350-7
; Sequence 7, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid
US-09-998-350-6

```

```

Query Match          97.8%; Score 44; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Qy      2 LYENVGMVY 9
        |||||

```

```
Db      2 LYENVGMVY 9

```

```

; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; TITLE OF INVENTION: SYNTHESIS AND USE
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Glu, which is di-tert-butoxy-gamma-carboxy-L-glutam
; OTHER INFORMATION: ic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: Tyr at position 3 is modified to Tyr(tBu), which is tert-butyl-ty
; OTHER INFORMATION: rosin
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
; OTHER INFORMATION: glutamic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is is trytyl-asp
; OTHER INFORMATION: argine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 is modified to Tyr(tBu), which is tert-butyl-ty
; OTHER INFORMATION: rosin
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys at position 10 is modified to Cys(Trt), which is trytyl-cyste
; OTHER INFORMATION: ine, and Cys(Trt) is connected to a resin
US-09-998-350-6

```

```

Query Match          97.8%; Score 44; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Qy      2 LYENVGMVY 9
        |||||

```

```
Db      2 LYENVGMVY 9

```

RESULT 8

```

US-09-998-350-8
; Sequence 8, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys at position 10 is modified to Cys(Trt), which is trytyl-cyste
; OTHER INFORMATION: ine, and Cys(Trt) is connected to a resin
US-09-998-350-6

```

```

Query Match          97.8%; Score 44; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Qy      2 LYENVGMVY 9
        |||||

```

```
Db      2 LYENVGMVY 9

```


; NAME/KEY: misc feature
 ; LOCATION: (1)..(1)
 ; OTHER INFORMATION: Glu at position 1 is modified to Glu(OtBu), which is tert-butoxy-
 ; OTHER INFORMATION: glutamic acid
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (4)..(4)
 ; OTHER INFORMATION: Glu at position 4 is modified to Glu(OtBu), which is tert-butoxy-
 ; OTHER INFORMATION: glutamic acid
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (5)..(5)
 ; OTHER INFORMATION: Asn at position 5 is modified to Asn(Trt), which is trytyl-aspara-
 ; OTHER INFORMATION: gine
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (9)..(9)
 ; OTHER INFORMATION: Tyr at position 9 is modified to Tyr(OtBu), which is tert-butoxy-
 ; OTHER INFORMATION: tyrosine
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (10)..(10)
 ; OTHER INFORMATION: Xaa = Adi(OAl), which is allyloxy-alpha-amino-adipic acid
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (10)..(10)
 ; OTHER INFORMATION: Xaa is an amide, i.e., C(O)NH
 ; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid
 US-09-998-350-14

Query Match 97.8%; Score 44; DB 10; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.032;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LYENVGMY 9
 |||||
 Db 2 LYENVGMY 9

RESULT 10

US-10-013-815-32
 ; Sequence 32, Application US/10013815
 ; Publication No. US20030105000A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Petro, Stephanie
 ; APPLICANT: Krug, David
 ; APPLICANT: Oligino, Lyn
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING GRB7
 ; FILE REFERENCE: V0139/7048 (HCL/MAT)
 ; CURRENT APPLICATION NUMBER: US/10/013,815
 ; CURRENT FILING DATE: 2001-11-05
 ; PRIOR APPLICATION NUMBER: US 60/245,755
 ; PRIOR FILING DATE: 2000-11-03
 ; NUMBER OF SEQ ID NOS: 194
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 32
 ; LENGTH: 11
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: No. US20030105000A1-phosphorylated peptide with YEN motif
 US-10-013-815-32

Query Match 97.8%; Score 44; DB 14; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LYENVGMY 9
 |||||
 Db 3 LYENVGMY 10

RESULT 11

US-09-998-350-18

; Sequence 18, Application US/09998350
 ; Publication No. US20030078368A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roller, Peter P
 ; APPLICANT: Long, Ya-Qiu
 ; APPLICANT: Lung, Feng-Di T
 ; APPLICANT: King, Richter C
 ; APPLICANT: Yang, Dajun
 ; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
 ; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND N
 ; TITLE OF INVENTION: SYNTHESIS AND USE
 ; FILE REFERENCE: 214683
 ; CURRENT APPLICATION NUMBER: US/09/998,350
 ; CURRENT FILING DATE: 2002-12-09
 ; PRIOR APPLICATION NUMBER: PCT/US00/15201
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: 60/137,187
 ; PRIOR FILING DATE: 1999-06-02
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 18
 ; LENGTH: 26
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(1)
 ; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid
 US-09-998-350-18

Query Match 97.8%; Score 44; DB 10; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.088;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LYENVGMY 9
 |||||
 Db 2 LYENVGMY 9

RESULT 12

US-09-998-350-19
 ; Sequence 19, Application US/09998350
 ; Publication No. US20030078368A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roller, Peter P
 ; APPLICANT: Long, Ya-Qiu
 ; APPLICANT: Lung, Feng-Di T
 ; APPLICANT: King, Richter C
 ; APPLICANT: Yang, Dajun
 ; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
 ; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND N
 ; TITLE OF INVENTION: SYNTHESIS AND USE
 ; FILE REFERENCE: 214683
 ; CURRENT APPLICATION NUMBER: US/09/998,350
 ; CURRENT FILING DATE: 2002-12-09
 ; PRIOR APPLICATION NUMBER: PCT/US00/15201
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: 60/137,187
 ; PRIOR FILING DATE: 1999-06-02
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 19
 ; LENGTH: 26
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(1)
 ; OTHER INFORMATION: Xaa = Glu, which is gamma-carboxy-L-glutamic acid

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa (Gla) has a CH2CO- group attached
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
US-09-998-350-19

```

```

Query Match      97.8%; Score 44; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.088;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY  2 LYENVGMV 9
    |||||
DB  2 LYENVGMV 9

```

```

RESULT 13
US-09-998-350-2
; Sequence 2, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic

```

```

; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa at position 1 is alpha-amino-adipic acid (Adi)
;
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa at position 4 is Glu or Adi
;
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Tyr at position 9 is an amide, i.e., C(O)NH
;
; NAME/KEY: misc_feature
; LOCATION: (1)..(9)
; OTHER INFORMATION: Xaa at position 1 and Tyr at position 9 are bridged together, making this
; OTHER INFORMATION: ing this peptide cyclic
US-09-998-350-2

```

```

Query Match      84.4%; Score 38; DB 10; Length 9;
Best Local Similarity 87.5%; Pred. No. 1.2e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY  2 LYENVGMV 9
    |||||
DB  2 LYENVGMV 9

```

```

RESULT 14
US-09-998-350-9
; Sequence 9, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: At position 1, Xaa = Adi, which is alpha-amino-adipic acid
;
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: At position 4, Xaa = Adi, which is alpha-amino-adipic acid
;
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
;
; NAME/KEY: misc_feature
; LOCATION: (1)..(10)
; OTHER INFORMATION: Xaa (Adi) at position 1 and Cys are bridged together, making this
; OTHER INFORMATION: peptide cyclic
US-09-998-350-9

```

```

Query Match      84.4%; Score 38; DB 10; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY  2 LYENVGMV 9
    |||||
DB  2 LYENVGMV 9

```

```

RESULT 15
US-09-998-350-10
; Sequence 10, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: At position 1, Xaa = Adi, which is alpha-amino-adipic acid
;
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: At position 4, Xaa = Adi, which is alpha-amino-adipic acid
;
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
;
; NAME/KEY: misc_feature
; LOCATION: (1)..(10)
; OTHER INFORMATION: Xaa (Adi) at position 1 and Cys are bridged together, making this
; OTHER INFORMATION: peptide cyclic
US-09-998-350-9

```

```

Query Match      84.4%; Score 38; DB 10; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY  2 LYENVGMV 9
    |||||
DB  2 LYENVGMV 9

```

```

RESULT 15
US-09-998-350-10
; Sequence 10, Application US/09998350
; Publication No. US20030078368A1
; GENERAL INFORMATION:
; APPLICANT: Roller, Peter P
; APPLICANT: Long, Ya-Qiu
; APPLICANT: Lung, Feng-Di T
; APPLICANT: King, Richter C
; APPLICANT: Yang, Dajun
; TITLE OF INVENTION: REDOX-STABLE, NON-PHOSPHORYLATED CYCLIC PEPTIDE INHIBITORS OF SH2
; TITLE OF INVENTION: BINDING TO TARGET PROTEIN, CONJUGATES THEREOF, COMPOSITIONS AND
; FILE REFERENCE: 214683
; CURRENT APPLICATION NUMBER: US/09/998,350
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: At position 1, Xaa = Adi, which is alpha-amino-adipic acid
;
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: At position 4, Xaa = Adi, which is alpha-amino-adipic acid
;
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Cys is an amide, i.e., C(O)NH
;
; NAME/KEY: misc_feature
; LOCATION: (1)..(10)
; OTHER INFORMATION: Xaa (Adi) at position 1 and Cys are bridged together, making this
; OTHER INFORMATION: peptide cyclic
US-09-998-350-9

```



```
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US00/15201
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,187
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Glu has a CH2CO- group attached
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: Xaa = Nle, which is norleucine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(10)
; OTHER INFORMATION: Glu and Cys are bridged together, making this peptide cyclic
US-09-998-350-10

Query Match      84.4%; Score 38; DB 10; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.5;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 LYENVGMY 9
Db 2 LYENVGXY 9
```

Search completed: July 20, 2004, 15:46:30
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 15:32:08 ; Search time 13 seconds
(without alignments)
36.049 Million cell updates/sec

Title: US-09-998-350-1
Perfect score: 45
Sequence: 1 XLYENVGMV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 38 | 84.4 | 565 | 1 PHR_YEAST | P05066 saccharomyc |
| 2 | 36 | 80.0 | 468 | 1 HEX_ADE31 | P36855 human adeno |
| 3 | 36 | 80.0 | 526 | 1 VGLG_SIGMA | P12647 sigma virus |
| 4 | 36 | 80.0 | 919 | 1 HEX_ADE12 | P19900 human adeno |
| 5 | 35 | 77.8 | 244 | 1 CYAH_MYRVE | P22143 myrothecium |
| 6 | 35 | 77.8 | 447 | 1 HEX_ADE04 | P36850 human adeno |
| 7 | 35 | 77.8 | 562 | 1 HEMA_IAPAP | P03451 influenza a |
| 8 | 34 | 75.6 | 1101 | 1 DIA2_HUMAN | O60879 homo sapien |
| 9 | 33 | 73.3 | 99 | 1 YLM3_CABEL | P34406 caenorhabdi |
| 10 | 33 | 73.3 | 312 | 1 FXXH_EAEIN | P44450 haemophilus |
| 11 | 33 | 73.3 | 754 | 1 RAD4_YEAST | P47336 saccharomyc |
| 12 | 32 | 71.1 | 306 | 1 PYRB_METUA | Q58976 methanococc |
| 13 | 32 | 71.1 | 313 | 1 CEO2_LACLA | P15244 lactococcus |
| 14 | 32 | 71.1 | 437 | 1 PAAK_ECOLI | P76085 escherichia |
| 15 | 32 | 71.1 | 512 | 1 VENV_THOVI | P28977 thogoto vir |
| 16 | 32 | 71.1 | 693 | 1 AGU1_SULSO | Q59645 sulfolobus |
| 17 | 32 | 71.1 | 1018 | 1 CONT_HUMAN | Q12860 homo sapien |
| 18 | 32 | 71.1 | 1020 | 1 CONT_MOUSE | P2960 mus musculu |
| 19 | 32 | 71.1 | 1021 | 1 CONT_RAT | Q63198 rattus norv |
| 20 | 31 | 68.9 | 221 | 1 Y805_METUA | Q58215 methanococ |
| 21 | 31 | 68.9 | 307 | 1 METF_STRLI | O54235 streptomyce |
| 22 | 31 | 68.9 | 405 | 1 C1WD_RAT | Q9890 rattus norv |
| 23 | 31 | 68.9 | 450 | 1 DCOR_CHICK | P27118 gallus gall |
| 24 | 31 | 68.9 | 455 | 1 DCOR_CRIGR | P14019 cricetus |
| 25 | 31 | 68.9 | 456 | 1 DCOR_XENLA | Q91854 xenopus lae |
| 26 | 31 | 68.9 | 460 | 1 DCOR_XENLA | P27120 xenopus lae |
| 27 | 31 | 68.9 | 461 | 1 DCOR_BOVIN | P27117 bos taurus |
| 28 | 31 | 68.9 | 461 | 1 DCOR_HUMAN | P11926 homo sapien |
| 29 | 31 | 68.9 | 461 | 1 DCOR_MOUSE | P00860 mus musculu |
| 30 | 31 | 68.9 | 461 | 1 DCOR_YUSPA | P27119 mus pahari |
| 31 | 31 | 68.9 | 461 | 1 DCOR_RAT | P09057 rattus norv |
| 32 | 31 | 68.9 | 519 | 1 ALGG_PSEBK | Q889c9 pseudomonas |
| 33 | 31 | 68.9 | 536 | 1 ALGG_PSEBM | Q887q3 pseudomonas |

ALIGNMENTS

RESULT 1

| PHR_YEAST | PHR_YEAST | STANDARD; | PRT; | 565 AA. |
|-----------|--|-----------|------|---------|
| ID | PHR_YEAST | STANDARD; | PRT; | 565 AA. |
| AC | P05066; | | | |
| DT | 13-AUG-1987 (Rel. 05, Created) | | | |
| DT | 13-AUG-1987 (Rel. 05, Last sequence update) | | | |
| DT | 28-FEB-2003 (Rel. 41, Last annotation update) | | | |
| DE | deoxyribodipyrimidine photolyase, mitochondrial precursor | | | |
| DE | (EC 4.1.99.3) (DNA photolyase) (Photoreactivating enzyme). | | | |
| GN | PHR1 OR YOR386W. | | | |
| OS | Saccharomyces cerevisiae (Baker's Yeast). | | | |
| OC | Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; | | | |
| OC | Saccharomycetales; Saccharomycetaceae; Saccharomycetes. | | | |
| OX | NCBI_TaxID=4932; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=86067229; PubMed=3906569; | | | |
| RA | Sancar G.B.; | | | |
| RT | "Sequence of the Saccharomyces cerevisiae PHR1 gene and homology of | | | |
| RT | the PHR1 photolyase to E. coli photolyase."; | | | |
| RL | Nucleic Acids Res. 13:8231-8246(1985). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=86083177; PubMed=3000886; | | | |
| RA | Yasui A., Langeveid S.A.; | | | |
| RT | "Homology between the photoreactivation genes of Saccharomycetes | | | |
| RT | cerevisiae and Escherichia coli."; | | | |
| RL | Gene 36:349-355(1985). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Delius H., Hebling U., Hofmann B.; | | | |
| RL | Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases. | | | |
| RN | [4] | | | |
| RP | REVIEW. | | | |
| RA | Sancar G.B., Sancar A.; | | | |
| RT | "Structure and function of DNA photolyases."; | | | |
| RT | Trends Biochem. Sci. 12:259-261(1987). | | | |
| RL | "FUNCTION: This enzyme catalyzes the light-dependent monomerization | | | |
| CC | (300-600 nm) of cyclobutyl pyrimidine dimers (in cis-syn | | | |
| CC | configuration), which are formed between adjacent bases on the | | | |
| CC | same DNA strand, upon exposure to ultraviolet radiation. | | | |
| CC | -/- CATALYTIC ACTIVITY: Cyclobutadipyrimidine (in DNA) = 2 pyrimidine | | | |
| CC | residues (in DNA). | | | |
| CC | -/- COFACTOR: Contains 2 chromophores: a reduced flavin (FADH2) and a | | | |
| CC | 5,10-methenyltetrahydrofolate. Both chromophores are bound by non- | | | |
| CC | covalent interactions. | | | |
| CC | -/- SUBCELLULAR LOCATION: Nuclear and mitochondrial. | | | |
| CC | -/- MISCELLANEOUS: This protein belongs to the "short wavelength-type | | | |
| CC | photolyases" with an absorption maximum at about 380 nm. | | | |
| CC | -/- MISCELLANEOUS: There are only 150-300 molecules of photolyase per | | | |
| CC | yeast cell. | | | |
| CC | -/- SIMILARITY: Belongs to the DNA photolyase class-1 family. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |

Q90687 gallus gall
Q06124 homo sapien
P41499 rattus norv
Q15546 homo sapien
P52891 saccharomyc
P16340 d trifuncti
P37297 saccharomyc
Q09246 caenorhabdi
Q45900 bacillus su
P59516 buchnera ap
P75519 mycoplasma
P32175 escherichia

34 31 68.9 593 1 PTNB_CHICK
35 31 68.9 593 1 PTNB_HUMAN
36 31 68.9 593 1 PTNB_RAT
37 31 68.9 671 1 RIKI_HUMAN
38 31 68.9 726 1 NUS4_YEAST
39 31 68.9 1364 1 PUR2_DROPS
40 31 68.9 1900 1 STT4_YEAST
41 30 66.7 177 1 YP98_CAEEL
42 30 66.7 178 1 YQAC_BACSU
43 30 66.7 201 1 LEUD_BUCBP
44 30 66.7 251 1 Y116_MYCPN
45 30 66.7 300 1 FDOH_ECOLI

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X03183; CAA26944.1; -
 EMBL: M1578; AAA34875.1; -
 EMBL: 275294; CAA99718.1; -
 PIR: S67298; S67298.
 GERMOnline: 143974; 1DNP.
 SGD: S000591.3; PH31.
 InterPro: IPR002081; DNA_photolyase_1.
 InterPro: IPR006050; DNA_photolyase_N.
 InterPro: IPR005101; FAD binding 7.
 InterPro: IPR006051; FAD binding_N.
 Pfam: PF00875; DNA_photolyase; 1.
 Pfam: PF03441; FAD binding 7; 1.
 PRINTS: PR00147; DNAPHOTLYASE.
 ProDom: PD004390; FAD binding_N; 1.
 PROSITE: PS00394; DNA_PHOTOLYASES_1; 1.
 PROSITE: PS00691; DNA_PHOTOLYASES_1_2; 1.
 Lyase; Chromophore; Flavoprotein; FAD; DNA repair; DNA-binding;
 Nuclear protein; Mitochondrion; Transit peptide.
 TRANSIT 1
 MITOCHONDRION.
 CHAIN ? 565
 DEOXYRIBODIPYRIMIDINE PHOTOLYASE.
 H-T-H MOTIF (POTENTIAL).
 V-> A (IN REF. 2).
 T-> S (IN REF. 2).
 S-> T (IN REF. 2).
 D-> S (IN REF. 2).
 S-> R (IN REF. 2).
 G-> E (IN REF. 2).
 E-> K (IN REF. 2).
 SEQUENCE 565 AA; 66274 MW; CD4FC3DA6128B97C CRC64;

Query Match 84.4%; Score 38; DB 1; Length 565;
 Best Local Similarity 75.0%; Pred. No. 3;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
 DB 87 LYDNVGLY 94

RESULT 2
 ID HEX_ADE31 STANDARD; PRT; 468 AA.
 AC F36855;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Hexon protein (Late protein 2) (Fragment).
 GN P1.
 OS Human adenovirus type 31.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=10529;
 RN [1]
 RP STRAIN=VRL 15/62; PubMed=8023012;
 RX MEDLINE=94294642; PubMed=8023012;
 RA Pring-Akerblom P., Adrian T.;
 RT "Type- and group-specific polymerase chain reaction for adenovirus detection.";
 RL Res. Virol. 145:25-35(1994).
 CC -!- FUNCTION: This protein is one of the structural proteins in the viral coat and is synthesized during late infection.
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X74661; CAA52725.1; -
 PIR: S37217; S37217.
 HSP: P03277; IDHX.
 InterPro: IPR000736; Adeno_hexon.
 Pfam: PF01065; Adeno_hexon; 1.
 ProDom: PD002815; Adeno_hexon; 1.
 KW Coat protein; Hexon protein; Late protein.
 FT NON_TER 1
 FT NON_TER 468 468
 SQ SEQUENCE 468 AA; 52100 MW; 8727BFA49179CE68 CRC64;

Query Match 80.0%; Score 36; DB 1; Length 468;
 Best Local Similarity 75.0%; Pred. No. 6.4;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
 DB 342 LYSNVGLY 349

RESULT 3
 ID VGLG_SIGMA STANDARD; PRT; 526 AA.
 AC P12647;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Spike glycoprotein precursor.
 GN G.
 OS Sigma virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; unclassified Rhabdoviridae.
 OX NCBI_TaxID=11301;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88034947; PubMed=2822842;
 RA Teninges D., Bras-Herreng F.;
 RT "Rhabdovirus sigma, the hereditary CO2 sensitivity agent of Drosophila: nucleotide sequence of a cDNA clone encoding the glycoprotein.";
 RT J. Gen. Virol. 68:2625-2638(1987).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X06171; CAA29536.1; -
 PIR: A27150; VGNVSG.
 FlyBase: FBgn0015809; Sigma-Virus[G].
 InterPro: IPR01903; Rhabd_glycop.
 Pfam: PF00974; Rhabdo_glycop; 1.
 KW Transmembrane; Envelope protein; Glycoprotein; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 526 SPIKE GLYCOPROTEIN.
 FT CARBOHYD 32 32 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 526 AA; 59010 MW; 335607C69249DD9D CRC64;

Query Match 80.0%; Score 36; DB 1; Length 526;
 Best Local Similarity 75.0%; Pred. No. 7.3;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
 DB 342 LYSNVGLY 349

```

Db      351 LYQSVGMV 358
RESULT 4
HEX_ADE12
ID HEX_ADE12 STANDARD; PRT; 919 AA.
AC P19900;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hexon protein (Late protein 2).
GN P11
OS Human adenovirus type 12.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94076430; PubMed=8254750;
RA Sprengel J., Schmitz B., Heuss-Neitzel D., Zock C., Doerfler W.;
RT "Nucleotide sequence of human adenovirus type 12 DNA: comparative functional analysis.";
RL J. Virol. 68:379-389 (1994).
RN [2]
RP SEQUENCE OF 888-919 FROM N.A.
RX STRAIN=Pereira 1131;
RX MEDLINE=88303354; PubMed=3043380;
RA Weber J.M., Houde A.;
RT "The primary structure of human adenovirus type 12 protease.";
RL Nucleic Acids Res. 16:7195-7195 (1988).
CC -!- FUNCTION: This protein is one of the structural proteins in the viral coat and is synthesized during late infection.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X73487; CAA51891.1; -
DR EMBL; X07655; CAA30501.1; -
DR EMBL; X07655; CAB37192.1; -
DR PIR; S01730; S01730.
DR PIR; S33942; S33942.
DR HSP; P03277; 1DHX.
DR InterPro; IPR000736; Adeno_hexon.
DR Pfam; PF01065; Adeno_hexon; 1.
DR ProDom; PD002815; Adeno_hexon; 1.
KW Coat protein; Hexon protein; Late protein.
SQ SEQUENCE 919 AA; 103039 MW; B37167885A516288 CRC64;

Query Match 80.0%; Score 36; DB 1; Length 919;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYQSVGMV 9
|||
DB 440 LYSNVGLY 447

RESULT 5
CYAH MYRVE
ID CYAH MYRVE STANDARD; PRT; 244 AA.
AC P22143;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-1998 (Rel. 36, Last annotation update)
DE Cyanamide hydratase (EC 4.2.1.69) (Urea hydro-lyase).
GN CAH.
OS Myrothecium verrucaria.

Query Match 77.8%; Score 35; DB 1; Length 244;
Best Local Similarity 75.0%; Pred. No. 5.3;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYQSVGMV 9
|||
DB 170 LYDNVGMV 177

RESULT 6
HEX_ADE04
ID HEX_ADE04 STANDARD; PRT; 447 AA.
AC F36850;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hexon protein (Late protein 2) (Fragment).
GN P11.
OS Human adenovirus type 4.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Isolate RJ-67;
RX MEDLINE=95407102; PubMed=7676636;
RA Pring-Akerblom P., Trijssenaar J., Adrian T.;
RT "Sequence characterization and comparison of human adenovirus subgenus B and E hexons.";
RL Virology 212:232-236 (1995).
CC -!- FUNCTION: This protein is one of the structural proteins in the viral coat and is synthesized during late infection.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Myrothecium.
OX NCBI_TaxID=5532;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=DSM 2087;
RX MEDLINE=91239547; PubMed=2034671;
RA Maier-Greiner U.M., Obermaier-Skrobranek B.M.M., Estermaier L.M.,
RA Kammerloher W., Freund C., Wueffling C., Burkert U.I., Matern D.H.,
RA Bremer M., Eulitz M., Kuefrevioglu O.I., Hartmann G.R.;
RT "Isolation and properties of a nitrile hydratase from the soil fungus Myrothecium verrucaria that is highly specific for the fertilizer cyanamide and cloning of its gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:4260-4264 (1991).
CC -!- FUNCTION: When used as herbicide in agriculture, cyanamide can be transformed, after sowing, in soil fertilizing ammonia by the combined action of M. verrucaria cyanamide hydratase and urease.
CC -!- CATALYTIC ACTIVITY: Urea = cyanamide + H(2)O.
CC -!- COFACTOR: Zinc.
CC -!- SUBUNIT: Homohexamer.
CC -!- MISCELLANEOUS: This enzyme is highly specific for cyanamide.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59078; AAA33429.1; -
DR PIR; A39365; A39365.
DR InterPro; IPR006674; HD.
DR InterPro; IPR003607; Met_phosphohydro.
DR Pfam; PF01966; HD; 1.
DR SMART; SM00471; HDC; 1.
KW Lyase; Zinc.
SQ SEQUENCE 244 AA; 26966 MW; 80F0A11F30E31CE2 CRC64;

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; X76550; CAA54052.1; --
 DR PIR; S39296; S39296.
 DR HSSP; P03277; LDHX.
 DR InterPro; IPR000736; Adeno_hexon.
 DR Pfam; PF01065; Adeno_hexon; 1.
 DR ProDom; PD002815; Adeno_hexon; 1.
 KW Coat protein; Hexon protein; Late protein.
 FT NON_TER 1
 FT NON_TER 447
 FT NON_TER 447
 SQ SEQUENCE 447 AA; 49553 MW; A7AE1977F707BD4D CRC64;

Query Match 77.8%; Score 35; DB 1; Length 447;
 Best Local Similarity 75.0%; Pred. No. 10;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
 |||:|:|:|
 Db 356 LYANVGLY 363

RESULT 7
 HEMA_IJAP
 ID HEMA_IJAP STANDARD; PRT; 562 AA.
 AC P03451;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hemagglutinin precursor [Contains: Hemagglutinin HA1 chain;
 DE Hemagglutinin HA2 chain].
 GN HA.
 OS Influenza A virus (strain A/Japan/305/57).
 OS Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 OC Influenza A viruses; Influenzavirus A.
 OX NCBI_TaxID=11421;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81030852; PubMed=7421990;
 RA Gething M.-J., Eys J., Skehel J.J., Waterfield M.;
 RT "Cloning and DNA sequence of double-stranded copies of haemagglutinin
 RT genes from H2 and H3 strains elucidates antigenic shift and drift in
 RT human influenza virus.";
 RL Nature 287:301-306(1980).
 CC -!- FUNCTION: Hemagglutinin is responsible for attaching the virus to
 CC cell receptors and for initiating infection.
 CC -!- SUBUNIT: Homotrimer. Each of the monomers is formed by two chains
 CC (HA1 and HA2) linked by a disulfide bond.
 CC -!- SIMILARITY: Belongs to the influenza viruses hemagglutinin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; J02127; AAA43185.1; --
 DR PIR; A04062; HMIV2.
 DR HSSP; P03437; IHFM.
 DR InterPro; IPR008980; Capsid_hemag.
 DR Pfam; PF00509; Hemagglutinin; 1.
 DR PRINTS; PR00329; HEMAGGLUTN12.
 DR ProDom; PD000225; Hemagglutn; 1.
 KW Envelope protein; Hemagglutinin; Glycoprotein; Signal.
 FT SIGNAL 1 15

FT CHAIN 16 339 HEMAGGLUTININ HA1 CHAIN.
 FT CHAIN 341 562 HEMAGGLUTININ HA2 CHAIN.
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 562 AA; 63118 MW; 6B7FD0C038993630 CRC64;

Query Match 77.8%; Score 35; DB 1; Length 562;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
 |||:|:|:|
 Db 204 LYQNVGTY 211

RESULT 8
 DIA2_HUMAN
 ID DIA2_HUMAN STANDARD; PRT; 1101 AA.
 AC O60879; O60878; Q9UUL2;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Diaphanous protein homolog 2 (Diaphanous-related formin 2) (DRP2).
 GN DIA2 OR DIA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE=98163437; PubMed=9497258;
 RA Bione S., Sala C., Manzini C., Arrigo G., Zuffardi O., Banfi S.,
 RA Borsani G., Jonveaux P., Philippe C., Zuccotti M., Sallabio A.,
 RA Toniolo D.;
 RT "A human homologue of the Drosophila melanogaster diaphanous gene is
 RT disrupted in a patient with premature ovarian failure: evidence for
 RT conserved function in oogenesis and implications for human
 RT sterility.";
 RL Am. J. Hum. Genet. 62:533-541(1998).
 [2]
 RN SEQUENCE OF 685-906 FROM N.A.
 RA Heath P.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: May be involved in oogenesis.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=DIA-156;
 CC IsoId=O60879-1; Sequences=Displayed;
 CC Name=DIA-12C;
 CC IsoId=O60879-2; Sequences=VSP 001573;
 CC TISSUE SPECIFICITY: Expressed in testis, ovary, small intestine,
 CC prostate, lung, liver, kidney, leukocytes.
 CC -!- DEVELOPMENTAL STAGE: Expressed from E16 in ovary and testis and
 CC during P6-P16 during differentiation of ovarian follicles.
 CC -!- DOMAIN: DRPs are regulated by intramolecular GBD-DAD binding where
 CC Rho-GTP activates the DRPs by disrupting the GBD-DAD interaction
 CC (By similarity).
 CC -!- DISPAR: Defects in DIA2 are a cause of premature ovarian
 CC failure (POF) (MIM:311360).
 CC -!- SIMILARITY: Contains 1 GTPase-binding (GBD) domain.
 CC -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
 CC -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
 CC -!- SIMILARITY: Contains 1 Formin homology 3 (FH3) domain.
 CC -!- SIMILARITY: Contains 1 DRF autoregulatory (DAD) domain.
 CC -!- SIMILARITY: Belongs to the formin homology family. Diaphanous
 CC subfamily.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; Y15909; CAA75870.1; -.
DR EMBL; Y15908; CAA75869.1; -.
DR EMBL; ALQ31053; CAB39108.1; -.
DR Genew; HGNC:2877; DIAPH2.
DR MIM; 300108; -.
DR MIM; 311360; -.
DR GO; GO:0005102; F:receptor binding; TAS.
DR GO; GO:0016288; P:cytokinesis; TAS.
DR GO; GO:0007292; P:female gamete generation; TAS.
DR InterPro; IPR003104; FH2.
DR Pfam; PF02181; FH2; 1.
DR SMART; SM00498; FH2; 1.
DR KW Alternative splicing; Coiled coil; Repeat.
FT DOMAIN 86 285
FT DOMAIN 184 482
FT DOMAIN 366 418
FT DOMAIN 487 547
FT DOMAIN 549 623
FT DOMAIN 628 1071
FT DOMAIN 903 1053
FT DOMAIN 1054 1068
FT DOMAIN 1072 1075
FT DOMAIN 257 260
FT DOMAIN 543 546
FT DOMAIN 562 572
FT DOMAIN 576 585
FT DOMAIN 591 597
FT DOMAIN 603 608
FT DOMAIN 613 616
FT DOMAIN 1038 1041
FT VARSPLIC 1081 1101
SQ SEQUENCE 1101 AA; 125568 MW; 399FIC292D79188B CRC64;
Query Match 75.6%; Score 34; DB 1; Length 1101;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 2 LYENVGMV 9
Db 971 LYENLGEY 978
RESULT 9
YLW3 CABEL
ID YLW3 CABEL STANDARD; PRT; 99 AA.
AC P34406;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein F22B7.3 in chromosome III.
GN F22B7.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodirinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=79063198;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,

```

```

RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P., Watson A., Weinstock L., Wilkinson-Sproat J.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; L12018; AAA65463.1; -.
DR PIR; S44632; S44632.
DR WormPep; F22B7.3; CE00156.
KW Hypothetical protein.
SQ SEQUENCE 99 AA; 11665 MW; 78FC94BDB3C8B585 CRC64;
Query Match 73.3%; Score 33; DB 1; Length 99;
Best Local Similarity 71.4%; Pred. No. 5.4;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 3 YENVGMV 9
Db 21 YENLGMF 27
RESULT 10
FDXH EAEIN
ID FDXH EAEIN STANDARD; PRT; 312 AA.
AC P44450;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Formate dehydrogenase, iron-sulfur subunit (Formate dehydrogenase beta
DE subunit) (FDH beta subunit).
GN FDXH OR H10007.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=96350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RA Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: ALLOWS TO USE FORMATE AS MAJOR ELECTRON DONOR DURING
CC ANAEROBIC RESPIRATION. THE BETA CHAIN IS AN ELECTRON TRANSFER UNIT
CC CONTAINING 4 CYSTEINE CLUSTERS INVOLVED IN THE FORMATION OF IRON-
CC SULFUR CENTRES. ELECTRONS ARE TRANSFERRED FROM THE GAMMA CHAIN TO
CC THE MOLYBDENUM COPACITOR OF THE ALPHA SUBUNIT (BY SIMILARITY).
CC -!- SUBUNIT: FORMATE DEHYDROGENASE IS A MEMBRANE-BOUND COMPLEX, FORMED
CC BY SUBUNITS ALPHA, BETA AND GAMMA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: ORTHOLOG OF BOTH E.COLI FDXH AND FDOH.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32686; AAC21685.1; -.
DR PIR; A64042; A64042.
DR HSSP; P00193; 1DUR.
DR TIGR; H10007; -.
DR InterPro; IPR001450; 4Fe4S ferredoxin.
DR TIGRFAMs; TIGR01582; FDH_beta.
DR Pfam; PF00037; fer4; 1.
DR InterPro; IPR006470; FDH_beta.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
DR Electron transport; 4Fe-4S; Iron-sulfur; Transmembrane;
KW Complete proteome.
FT METAL 44 44 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 47 47 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 50 50 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 54 54 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 106 106 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 109 109 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 114 114 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 118 118 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 139 139 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 142 142 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 145 145 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 149 149 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 166 166 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
FT METAL 169 169 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
FT METAL 181 181 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
FT METAL 185 185 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 312 AA; 34068 MW; AA49DD3C7064866 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 312;
Best Local Similarity 71.4%; Pred. No. 18;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 YENVGMY 9
| | | | |
Db 214 YENAGLY 220

RESULT 11
RAD4_YEAST STANDARD; PRT; 754 AA.
AC P14736;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA repair protein RAD4.
GN RAD4 OR YER162C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232744; PubMed=3073107;
RA Gietz R.D., Prakash S.;
RT "Cloning and nucleotide sequence analysis of the Saccharomyces
RT cerevisiae RAD4 gene required for excision repair of UV-damaged
RT DNA."
RL Gene 74:535-541 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89197751; PubMed=2649477;
RA Couto L.B., Friedberg E.C.;
RT "Nucleotide sequence of the wild-type RAD4 gene of Saccharomyces

```

```

RT cerevisiae and characterization of mutant rad4 alleles."
RL J. Bacteriol. 171:1862-1869 (1989).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / A9772;
RX MEDLINE=97313264; PubMed=9169868;
RX Dietrich F.S., Mulligan J.T., Brennan T., Carpenter J., Chen E.,
RX Araujo R., Aviles E., Berno A., Guzman M., Hartzell G.,
RX Cherry J.M., Chung E., Duncan M., Kayser A., Komp C., Lashkari D., Lew H.,
RX Hunnicke-Smith S., Hyman R.W., Namath A., Norgren R., Oefner P.,
RX Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RX Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RX Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V."
RL Nature 387:78-81 (1997).
CC -1- FUNCTION: Involved in nucleotide excision repair of DNA damaged
CC with UV light, bulky adducts, or cross-linking agents.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the XPC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26050; AAA34944.1; -.
DR EMBL; M24928; AAA34945.1; -.
DR EMBL; U18917; AAB64689.1; -.
DR PIR; S30814; DDBD4.
DR Germline; L39239; -.
DR SGD; S0000364; RAD4.
DR GO; GO:0000111; C:nucleotide excision repair factor 2 complex; IDA.
DR GO; GO:000108; C:repirosome; IDA.
DR GO; GO:0003684; F:damaged DNA binding; IDA.
DR InterPro; IPR004583; Rad4.
DR Pfam; PF03835; Rad4; 1.
DR TIGRFAMs; TIGR00605; rad4; 1.
KW DNA repair; DNA-binding; Nuclear protein.
FT DNA BIND 250 269 POTENTIAL.
FT CONFLICT 223 225 VGI -> EGL (IN REF. 3).
SQ SEQUENCE 754 AA; 87174 MW; 788C146DC4BD2BF8 CRC64;

Query Match 73.3%; Score 33; DB 1; Length 754;
Best Local Similarity 71.4%; Pred. No. 45;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 YENVGMY 9
| | | | |
Db 220 YENVGMY 226

RESULT 12
PYRB_METJA STANDARD; PRT; 306 AA.
AC G58976;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Aspartate carbamoyltransferase (EC 2.1.3.2) (Aspartate
DE transcarbamylase) (ATCase).
GN PYRB OR MJ1581.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96337999; PubMed=8688097;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,

```


RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weissflog K.G., Merrick J.M., Glöck A.,
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.,
 FT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RL jannaschii.",
 RL Science 273:1058-1073 (1996).
 RN [2]
 RN CHARACTERIZATION.
 RP MEDLINE=20283607; PubMed=10748118;
 RX Hack E.S., Vorobyova T., Sakash J.B., West J.M., Maccl C.P., Herve G.,
 RA Williams M.K., Kantrowitz E.R.;
 RA "Characterization of the aspartate transcarbamoylase from
 RT Methanococcus jannaschii.",
 RL Acta Crystallogr. D 56:1061-1063 (2000).
 CC [3]
 CC CRYSTALLIZATION, AND X-RAY CRYSTALLOGRAPHY.
 RN MEDLINE=20402716; PubMed=10944354;
 RX Vitali J., Vorobyova T., Webster G., Kantrowitz E.R.;
 RA "Crystallization and structure determination of the catalytic trimer
 RT of Methanococcus jannaschii aspartate transcarbamoylase.",
 RL Acta Crystallogr. D 56:1061-1063 (2000).
 CC [4]
 CC CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate = phosphate
 + N-carbamoyl-L-aspartate.
 CC [5]
 CC PATHWAY: Pyrimidine biosynthesis; second step.
 CC [6]
 CC SUBUNIT: HETERODECAMER (23:3R2); AND SIX CATALYTIC PYRIB CHAINS
 CC ORGANIZED AS TWO TRIMERS (C3); AND SIX REGULATORY PYRI CHAINS
 CC ORGANIZED AS THREE DIMERS (R2).
 CC [7]
 CC SIMILARITY: Belongs to the ATCase/OTCase family.
 CC [8]
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----
 DR EMBL; U67598; AAB95601.1; --
 DR PIR; D64497; D64497.
 DR HSP; P00479; 3CSU.
 DR TIGR; MJ1581; --
 DR HAMAP; MF 00001; -- 1.
 DR InterPro; IPR006130; Asp/Orn COTranf.
 DR InterPro; IPR002082; Asp.carbMltransf.
 DR InterPro; IPR006131; OTCase_O.
 DR InterPro; IPR006132; OTCase_P.
 DR Pfam; PF00185; OTCase; 1.
 DR Pfam; PF02729; OTCase N; 1.
 DR PRINTS; PR00100; AOTCASE.
 DR TIGRFAMs; TIGR00670; asp carb tr; 1.
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE, 1.
 KW Pyrimidine biosynthesis; Transferrase; Complete proteome.
 SQ SEQUENCE 306 AA; 35159 MW; CBDC31FC450CEFeA CRC64;

 Query Match 71.1%; Score 32; DB 1; Length 306;
 Best Local Similarity 75.0%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 LYENVGMY 9
 |:|:|:|:|
 Db 175 LFENVEMY 182

 RESULT 13
 CE02_LACIA STANDARD; PRT; 313 AA.
 ID_CFO2_LACIA
 AC P15244;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE N(5)-(carboxyethyl)ornithine synthase (EC 1.5.1.24) (N(5)-(L-1-
 DE carboxyethyl)-L-ornithine:NADP(+) oxidoreductase) (CEOS).
 GN CEO.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1350;
 RN [1]
 RN SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-15.
 RP STRAIN=K1-23; TRANSPOSON=Th5306;
 RC MEDLINE=95263576; PubMed=7744873;
 RX Donkersloot J.A., Thompson J.;
 RA "Cloning, expression, sequence analysis, and site-directed
 RT mutagenesis of the Th5306-encoded N5-(carboxyethyl)ornithine synthase
 RT from Lactococcus lactis K1.",
 RL J. Biol. Chem. 270:12226-12234 (1995).
 RN [2]
 RN SEQUENCE OF 1-37.
 RP STRAIN=K1;
 RX MEDLINE=89255467; PubMed=2498334;
 RA Thompson J.;
 RT "N5-(L-1-carboxyethyl)-L-ornithine:NADP+ oxidoreductase from
 RT Streptococcus lactis. Purification and partial characterization.",
 RL J. Biol. Chem. 264:9592-9601 (1989).
 RN [3]
 RN SEQUENCE OF 256-263, AND CHARACTERIZATION.
 RP STRAIN=K1;
 RX MEDLINE=20014035; PubMed=10548058;
 RA Sackett D.L., Ruvinov S.B., Thompson J.;
 RT "N5-(L-1-carboxyethyl)-L-ornithine synthase: physical and spectral
 RT characterization of the enzyme and its unusual low pKa fluorescent
 RT tyrosine residues.",
 RL Protein Sci. 8:2121-2129 (1999).
 RN [4]
 RN FOLDING STUDIES.
 RP STRAIN=K1;
 RX MEDLINE=99456521; PubMed=10525296;
 RA Ruvinov S.B., Thompson J., Sackett D.L., Ginsburg A.;
 RT "Tetraneric N(5)-(L-1-carboxyethyl)-L-ornithine synthase: guanidine.
 RT HCl-induced unfolding and a low temperature requirement for
 RT refolding",
 RL Arch. Biochem. Biophys. 371:115-123 (1999).
 CC [5]
 CC CATALYTIC ACTIVITY: N(5)-(L-1-carboxyethyl)-L-ornithine + NADP(+) +
 CC H(2)O = L-ornithine + pyruvate + NADPH.
 CC [6]
 CC SUBUNIT: Homotetramer.
 CC [7]
 CC MASS SPECTROMETRY: MW=35.355; METHOD=MALDI.
 CC [8]
 CC MISCELLANEOUS: In the reverse direction L-lysine can act instead
 CC of L-ornithine, more slowly, yielding N(6)-(L-1-carboxyethyl)-L-
 CC lysine.
 CC [9]
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).
 CC -----
 DR EMBL; U23376; AAA86385.1; --
 DR PIR; A57499; A57499.
 DR InterPro; IPR007698; AlaDh_PNT_C.
 DR InterPro; IPR007886; AlaDh_PNT_N.
 DR Pfam; PF01262; AlaDh_PNT_C7_1.
 DR Pfam; PF05222; AlaDh_PNT_N; 1.
 KW Oxidoreductase; NADP.
 FT BIND 171 176 NADPH (POTENTIAL).
 FT MUTAGEN 15 15 R->K; LOSS OF ACTIVITY.
 SQ SEQUENCE 313 AA; 35323 MW; B17FE0F477113C77 CRC64;

 Query Match 71.1%; Score 32; DB 1; Length 313;
 Best Local Similarity 62.5%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 15:36:14 ; Search time 16 Seconds
(without alignments)
54.108 Million cell updates/sec

Title: US-09-998-350-1

Perfect score: 45

Sequence: 1 XLYENVGY 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|------------------------|
| 1 | 38 | 84.4 | 565 | 2 S67298 | deoxyribodipyrimidine |
| 2 | 36 | 80.0 | 468 | 2 S37217 | hexon protein - hu |
| 3 | 36 | 80.0 | 526 | 1 VGVNSG | spike glycoprotein |
| 4 | 36 | 80.0 | 919 | 2 S33942 | hexon protein - hu |
| 5 | 35 | 77.8 | 20 | 2 P10161 | hemagglutinin - In |
| 6 | 35 | 77.8 | 244 | 2 A39365 | cyanamide hydratase |
| 7 | 35 | 77.8 | 447 | 2 S39296 | hexon protein - hu |
| 8 | 35 | 77.8 | 448 | 1 F70190 | probable diphosphatase |
| 9 | 35 | 77.8 | 562 | 1 HMIV2 | hemagglutinin prec |
| 10 | 35 | 77.8 | 936 | 2 S57637 | hexon protein - hu |
| 11 | 34 | 75.6 | 29 | 2 H81136 | hypothetical prote |
| 12 | 34 | 75.6 | 34 | 2 H81883 | hypothetical prote |
| 13 | 34 | 75.6 | 150 | 2 A55883 | actin-filament-ass |
| 14 | 33 | 73.3 | 99 | 2 S44632 | 122b7.3 protein - |
| 15 | 33 | 73.3 | 309 | 2 F83044 | nitrate-inducible |
| 16 | 33 | 73.3 | 312 | 2 A64042 | formate dehydrogen |
| 17 | 33 | 73.3 | 332 | 2 T33774 | hypothetical prote |
| 18 | 33 | 73.3 | 439 | 2 G88103 | protein W10G11.17 |
| 19 | 33 | 73.3 | 511 | 2 A99574 | ABC transporter at |
| 20 | 33 | 73.3 | 1249 | 2 A56511 | myosin I myoA - Em |
| 21 | 32 | 71.1 | 149 | 2 S67188 | hypothetical prote |
| 22 | 32 | 71.1 | 306 | 2 D64497 | aspartate carboxyl |
| 23 | 32 | 71.1 | 313 | 2 A57499 | N5-(carboxyethyl)O |
| 24 | 32 | 71.1 | 352 | 2 D72264 | hypothetical prote |
| 25 | 32 | 71.1 | 354 | 2 E97128 | magnesium and cova |
| 26 | 32 | 71.1 | 389 | 2 B81380 | hypothetical prote |
| 27 | 32 | 71.1 | 434 | 2 S50865 | avermectin-sensiti |
| 28 | 32 | 71.1 | 437 | 2 A64891 | coenzyme F390 synt |
| 29 | 32 | 71.1 | 512 | 1 VGIVTH | envelope glycoprot |

| | | | | | |
|----|----|------|------|----------|--------------------|
| 30 | 32 | 71.1 | 591 | 2 G95899 | asparagine synthas |
| 31 | 32 | 71.1 | 661 | 2 S49901 | coat protein gpl - |
| 32 | 32 | 71.1 | 688 | 2 T33708 | hypothetical prote |
| 33 | 32 | 71.1 | 693 | 2 H90486 | alpha-glucosidase |
| 34 | 32 | 71.1 | 700 | 2 T20550 | hypothetical prote |
| 35 | 32 | 71.1 | 739 | 2 A11876 | hypothetical prote |
| 36 | 32 | 71.1 | 852 | 2 T33824 | hypothetical prote |
| 37 | 32 | 71.1 | 1018 | 2 JG4211 | neural adhesion pr |
| 38 | 32 | 71.1 | 1018 | 2 A54744 | contactin 1 precur |
| 39 | 32 | 71.1 | 1020 | 2 S05944 | neuronal cell surf |
| 40 | 32 | 71.1 | 1021 | 2 A57112 | contactin precurs |
| 41 | 32 | 71.1 | 1181 | 2 D86157 | hypothetical prote |
| 42 | 31 | 68.9 | 221 | 2 E64400 | conserved hypothet |
| 43 | 31 | 68.9 | 224 | 2 H98847 | hypothetical prote |
| 44 | 31 | 68.9 | 231 | 2 H85138 | hypothetical prote |
| 45 | 31 | 68.9 | 234 | 2 S14237 | Ig kappa chain pre |

ALIGNMENTS

RESULT 1

S67298

deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) - Yeast (Saccharomyces cerevisiae)

N:Alternate names: protein O6771; protein YOR386W

C:Species: Saccharomyces cerevisiae

C>Date: 12-Jul-1996 #sequence revision 12-Jul-1996 #text_change 20-Jun-2000

C:Accession: S67298; A23964; A24046

R:Delius, H.; Hebling, U.; Hofmann, B.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67261

A:Accession: S67298

A:Molecule type: DNA

A:Residues: 1-565

A:Cross-references: EMBL:Z75294; NID:G1420830; PIDN:CAA99718.1; PID:G1420831; MIPS:YOR386

A:Experimental source: strain S288C

R:Yasui, A.; Langeveld, S.A.

Gene 36, 349-355, 1985

A:Title: Homology between the photoreactivation genes of Saccharomyces cerevisiae and Es

A:Reference number: A23964; MUID:86083177; PMID:3000886

A:Accession: A23964

A:Molecule type: DNA

A:Residues: 1-76, 'A', 78-164, 'S', 166-168, 'T', 170-199, 'S', 201-350, 'R', 352-364, 'E', 366-472, 'Nucleic Acids Res. 13, 8231-8246, 1985

R:Sancar, G.B.

A:Title: Sequence of the Saccharomyces cerevisiae PHR1 gene and homology of the PHR1 phot

A:Reference number: A24046; MUID:86067229; PMID:3906569

A:Accession: A24046

A:Molecule type: DNA

A:Residues: 1-585 <SAN>

A:Cross-references: EMBL:X03183; NID:G4175; PIDN:CAA26944.1; PID:G4176

C:Genetics:

A:Gene: SGD:PHR1

A:Cross-references: SGD:S0005913; MIPS:YOR386W

A:Map position: 15R

C:Superfamily: deoxyribodipyrimidine photo-lyase

C:Keywords: carbon-carbon lyase

Query Match 84.4%; Score 38; DB 2; Length 565;

Best Local Similarity 75.0%; Pred. No. 4.7;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGY 9

DB 87 LYDNVGLY 94

RESULT 2

S37217

hexon protein - human adenovirus 31 (fragment)

C:Species: Mastadenovirus h31 (human adenovirus 31)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999

| | |
|--|--|
| <p>C;Accession: S37217 R;Ping-Akerblom, P. submitted to the EMBL Data Library, September 1993 A;Reference number: S37213 A;Accession: S37217 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-468 <PRI> A;Cross-references: EMBL:X74661, NID:9402765; PIDN:CAA52725.1; PID:9402766 C;Superfamily: adenovirus hexon protein</p> | <p>Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;</p> <p>QY 2 LYENVGMY 9 : Db 440 LYSNVGLY 447</p> |
| <p>RESULT 5 hemagglutinin - Influenza H2N2 (fragment) C;Species: Influenza H2N2 C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-May-1997 C;Accession: PL0161 R;Sweetser, M.T.; Braciale, V.L.; Braciale, T.J. J. Exp. Med. 170, 1357-1368, 1989 A;Title: Class I major histocompatibility complex-restricted T lymphocyte recognition of A;Reference number: PL0161; MUID:90010790; PMID:2477491 A;Accession: PL0161 A;Molecule type: mRNA A;Residues: 1-20 <SWE> A;Experimental source: strain A/JAP/305/57 C;Comment: This protein plays a major role in initiation of infection and in the pathogen C;Superfamily: influenza virus hemagglutinin C;Keywords: hemagglutinin F;1-20/Region: immunodominant site recognized by T-lymphocytes</p> | <p>Query Match 77.8%; Score 35; DB 2; Length 20; Best Local Similarity 75.0%; Pred. No. 0.49; Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;</p> <p>QY 2 LYENVGMY 9 : Db 3 LYQNVGTY 10</p> |
| <p>RESULT 6 cyanamide hydratase (EC 4.2.1.69) - fungus (Myrothecium verrucaria) C;Species: Myrothecium verrucaria C;Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 15-Sep-2000 C;Accession: A39365 R;Maier-Greiner, U.H.; Obermaier-Skrobranek, B.M.M.; Estermaier, L.M.; Kammerloher, W.; R. Proc. Natl. Acad. Sci. U.S.A. 88, 4260-4264, 1991 A;Title: Isolation and properties of a nitrile hydratase from the soil fungus Myrothecium A;Reference number: A39365; MUID:91239547; PMID:2034671 A;Accession: A39365 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-244 <MAI> A;Cross-references: GB:M59078; NID:9168392; PIDN:AAA33429.1; PID:9168393 C;Superfamily: Saccharomyces cerevisiae hypothetical protein YFL061W C;Keywords: carbon-oxygen lyase; hydro-lyase</p> | <p>Query Match 77.8%; Score 35; DB 2; Length 244; Best Local Similarity 75.0%; Pred. No. 8.2; Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;</p> <p>QY 2 LYENVGMY 9 : Db 170 LYDNVGGY 177</p> |
| <p>RESULT 7 hexon protein - human adenovirus 4 C;Species: Mastadenovirus h4 (human adenovirus 4) C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999 C;Accession: S39296 R;Ping-Akerblom, P.; Adrian, T. submitted to the EMBL Data Library, November 1993 A;Reference number: S39296 A;Accession: S39296</p> | <p>Query Match 75.0%; Score 36; DB 2; Length 919; Best Local Similarity 75.0%; Pred. No. 22;</p> |
| <p>C;Accession: S37217 R;Ping-Akerblom, P. submitted to the EMBL Data Library, September 1993 A;Reference number: S37213 A;Accession: S37217 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-468 <PRI> A;Cross-references: EMBL:X74661, NID:9402765; PIDN:CAA52725.1; PID:9402766 C;Superfamily: adenovirus hexon protein</p> | <p>Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;</p> <p>QY 2 LYENVGMY 9 : Db 351 LYQSVGMY 358</p> |
| <p>RESULT 4 S33942 hexon protein - human adenovirus 12 N;Alternate names: late protein 2 C;Species: Mastadenovirus h12 (human adenovirus 12) C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999 C;Accession: S33942 R;Sprengel, J. submitted to the EMBL Data Library, June 1993 A;Reference number: S33928 A;Accession: S33942 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-919 <SPR> A;Cross-references: EMBL:X73487; NID:9313361; PIDN:CAA51891.1; PID:9313376 C;Superfamily: adenovirus hexon protein</p> | <p>Query Match 80.0%; Score 36; DB 2; Length 919; Best Local Similarity 75.0%; Pred. No. 22;</p> |
| <p>RESULT 3 VGVNSG spike glycoprotein G precursor - sigma virus C;Species: sigma virus A;Note: host Drosophila melanogaster C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999 C;Accession: A27150 R;Teninges, D.; Bras-Herrens, F. J. Gen. Virol. 68, 2625-2638, 1987 A;Title: Rhabdovirus sigma, the hereditary CO-2 sensitivity agent of Drosophila: nucleot A;Reference number: A27150; MUID:89034947; PMID:2822842 A;Accession: A27150 A;Molecule type: genomic RNA A;Residues: 1-526 <TEN> A;Cross-references: GB:X06171; NID:961818; PIDN:CAA29536.1; PID:961819 C;Genetics: A;Gene: G A;Cross-references: FlyBase:FBgn0015809 C;Superfamily: Rhabdovirus spike glycoprotein G C;Keywords: glycoprotein; spike protein; transmembrane protein F;1-17/Domain: signal sequence #status predicted <SIG> F;18-526/Product: spike glycoprotein G #status predicted <SGG> F;499-515/Domain: transmembrane #status predicted <TMN> F;52.445;459/Binding site: carbohydrate (Asn) (covalent) #status predicted</p> | <p>Query Match 80.0%; Score 36; DB 1; Length 526; Best Local Similarity 75.0%; Pred. No. 12; Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;</p> <p>QY 2 LYENVGMY 9 : Db 351 LYQSVGMY 358</p> |
| <p>RESULT 4 S33942 hexon protein - human adenovirus 12 N;Alternate names: late protein 2 C;Species: Mastadenovirus h12 (human adenovirus 12) C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999 C;Accession: S33942 R;Sprengel, J. submitted to the EMBL Data Library, June 1993 A;Reference number: S33928 A;Accession: S33942 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-919 <SPR> A;Cross-references: EMBL:X73487; NID:9313361; PIDN:CAA51891.1; PID:9313376 C;Superfamily: adenovirus hexon protein</p> | <p>Query Match 80.0%; Score 36; DB 2; Length 919; Best Local Similarity 75.0%; Pred. No. 22;</p> |

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <PR>
A:Cross-references: EMBL:X76550; NID:g434903; PIDN:CAA54052.1; PID:g434904
C:Superfamily: adenovirus hexon protein

Query Match 77.8%; Score 35; DB 2; Length 447;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
|||:|:|
Db 356 LYANVGLY 363

RESULT 8
F70190
Probable diphosphate-fructose-6-phosphate 1-phosphotransferase (EC 2.7.1.90) - Lyme disease
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 04-Aug-2003
C:Accession: F70190
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, S.D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, R.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: F70190
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-448 <LE>
A:Cross-references: GB:AE001172; GB:AE000783; NID:g2688654; PIDN:AAC67070.1; PID:g268865
A:Experimental source: strain B31
C:Superfamily: pyrophosphate-dependent phosphofructokinase, Eh/Ppi-PFK type; 6-phosphofructokinase
C:Keywords: phosphotransferase
F:82-398/Domain: 6-phosphofructokinase 1 homology <6PF>

Query Match 77.8%; Score 35; DB 1; Length 448;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
|||:|:|
Db 337 LYEDIGLY 344

RESULT 9
HM12
hemagglutinin precursor - influenza A virus (strain A/Japan/305/57 [H2])
C:Species: influenza A virus
A:Variety: strain A/Japan/305/57 [H2]
C>Date: 28-Feb-1981 #sequence_revision 28-Feb-1981 #text_change 16-Jul-1999
C:Accession: A04062; S12270
R:Gething, M.J.; Byr, J.; Skehel, J.; Waterfield, M. Nature 287, 301-306, 1980
A:Title: Cloning and DNA sequence of double-stranded copies of haemagglutinin genes from A/Japan/305/57 influenza virus
A:Reference number: A93233; MUID:81030852; PMID:7421990
A:Accession: A04062
A:Molecule type: mRNA
A:Residues: 1-562 <GET>
A:Cross-references: GB:J02127; NID:g324145; PIDN:AAA43185.1; PID:g324146
A:Experimental source: strain A/Japan/305/57 [H2]
R:Naeye, C.W.; Williams, D. EMBO J. 9, 3857-3866, 1990
A:Title: Fatty acids on the A/Japan/305/57 influenza virus hemagglutinin have a role in membrane fusion
A:Reference number: S12270; MUID:91065313; PMID:2249653
A:Accession: S12270
A:Molecule type: mRNA
A:Residues: 510-562 <NA>
A:Experimental source: strain A/Japan/305/57 (H2N2)
C:Superfamily: influenza virus hemagglutinin
C:Keywords: hemagglutinin; homotrimer; lipoprotein; thiolester bond

F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-329/Product: hemagglutinin chain HA1 #status predicted <HA1>
F:341-562/Product: hemagglutinin chain HA2 #status predicted <HA2>
F:551-558/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 77.8%; Score 35; DB 1; Length 562;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
|||:|:|
Db 204 LYQNVGT 211

RESULT 10
S57637
hexon protein - human adenovirus 4
C:Species: Mastadenovirus h4 (human adenovirus 4)
C>Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 26-Aug-1999
C:Accession: S57637
R:Pring-Akerblom, P.; Trijssenaar, J.; Adrian, T. submitted to the EMBL Data Library, February 1995
A:Reference number: S57637
A:Accession: S57637
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-936 <PRI>
A:Cross-references: EMBL:X84646; NID:g886486; PIDN:CAA59139.1; PID:g886487
C:Superfamily: adenovirus hexon protein

Query Match 77.8%; Score 35; DB 2; Length 936;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
|||:|:|
Db 457 LYANVGLY 464

RESULT 11
B81136
hypothetical protein NMB0968 [imported] - Neisseria meningitidis (strain MC58 serogroup I)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: B81136
R:Hettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Verma, M. Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: B81136
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-29 <NET>
A:Cross-references: GB:AE002448; GB:AE002098; NID:g7226204; PIDN:AAF41373.1; PID:g7226204
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0968

Query Match 75.6%; Score 34; DB 2; Length 29;
Best Local Similarity 62.5%; Pred. No. 1.2;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
|||:|:|
Db 22 LYKNLGLY 29

RESULT 12
H81883
hypothetical protein NMA165 [imported] - Neisseria meningitidis (strain Z2491 serogroup

C:Species: Neisseria meningitidis
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: H81883
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel,
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: H81883
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-34 <PAR>
 A:Cross-references: GB:AL162755; GB:AL157959; NID:G7379742; PID:G737965
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: NMA1165

Query Match 75.6%; Score 34; DB 2; Length 34;
 Best Local Similarity 62.5%; Pred. No. 1.5;
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
 |||:|:
 Db 27 LYKNLGLY 34

RESULT 13
 A5883
 actin-filament-associated protein 120k form - chicken' (fragment)
 C:Species: Gallus gallus (Chicken)
 C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
 C:Accession: A65883
 R:Flynn, D.C.; Koay, T.C.; Humphries, C.G.; Guappone, A.C.
 J. Biol. Chem. 270, 3894-3899, 1995
 A:Title: AFAP-120. A variant form of the src SH2/SH3-binding partner AFAP-110 is detected
 A:Reference number: A55883; MUID:95181352; PMID:7876134
 A:Accession: A5883
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-150 <FLY>
 A:Cross-references: GB:L20302

Query Match 75.6%; Score 34; DB 2; Length 150;
 Best Local Similarity 62.5%; Pred. No. 7.7;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMY 9
 |||:|:
 Db 52 LYDNAGLY 59

RESULT 14
 S44632
 F22B7.3 protein - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Sep-1997
 C:Accession: S44632
 R:Anderson, K.
 submitted to the EMBL Data Library, March 1993
 A:Description: Sequence of the C. elegans cosmid F22B7.
 A:Reference number: S44628
 A:Accession: S44632
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-99 <AND>
 A:Cross-references: EMBL:L12018; NID:G156298; PID:G156303
 C:Genetics:
 A:Introns: 25/1; 81/3

Query Match 73.3%; Score 33; DB 2; Length 99;
 Best Local Similarity 71.4%; Pred. No. 8;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 YENVGMY 9
 |||:|:
 Db 21 YENLGMF 27

RESULT 15
 F83044
 nitrate-inducible formate dehydrogenase, beta subunit PA4811 [imported] - Pseudomonas aer
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 02-Aug-2002
 C:Accession: F83044
 R:Stover, C.K.; Phan, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: F83044
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-309 <STO>
 A:Cross-references: GB:AE004894; GB:AE004091; NID:G9951076; PIDN:AAG08197.1; GSPDB:GH001;
 A:Experimental source: strain PAO1
 C:Superfamily: formate dehydrogenase, nitrate-inducible, beta chain; ferredoxin 2[4Fe-4S]

Query Match 73.3%; Score 33; DB 2; Length 309;
 Best Local Similarity 71.4%; Pred. No. 29;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 YENVGMY 9
 |||:|:
 Db 204 YENAGLY 210

Search completed: July 20, 2004, 15:45:07
 Job time : 16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 15:42:39 ; Search time 18 Seconds
(without alignments)
25.813 Million cell updates/sec

Title: US-09-998-350-1

Perfect score: 45

Sequence: 1 XLVENVGMV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 38 | 84.4 | 566 | 2 | US-08-272-255-8 |
| 2 | 38 | 84.4 | 566 | 5 | PCT-US95-08565-8 |
| 3 | 36 | 80.0 | 919 | 2 | US-08-788-874-4 |
| 4 | 35 | 77.8 | 19 | 1 | US-08-146-145-6 |
| 5 | 35 | 77.8 | 19 | 4 | US-08-376-343-3 |
| 6 | 35 | 77.8 | 20 | 2 | US-08-480-190-38 |
| 7 | 35 | 77.8 | 20 | 2 | US-08-488-379-38 |
| 8 | 35 | 77.8 | 20 | 4 | US-08-475-399A-38 |
| 9 | 35 | 77.8 | 20 | 5 | PCT-US93-07545-38 |
| 10 | 35 | 77.8 | 244 | 3 | US-08-003-287-6 |
| 11 | 35 | 77.8 | 244 | 3 | US-08-003-287-8 |
| 12 | 35 | 77.8 | 244 | 3 | US-09-518-988-2 |
| 13 | 34 | 75.6 | 362 | 2 | US-09-080-897-6 |
| 14 | 34 | 75.6 | 362 | 3 | US-09-323-735-6 |
| 15 | 33 | 73.3 | 310 | 4 | US-09-252-991A-27339 |
| 16 | 32 | 71.1 | 15 | 1 | US-08-176-500-31 |
| 17 | 32 | 71.1 | 15 | 1 | US-08-471-052A-31 |
| 18 | 32 | 71.1 | 15 | 1 | US-08-189-331-31 |
| 19 | 32 | 71.1 | 15 | 2 | US-08-471-839-31 |
| 20 | 32 | 71.1 | 15 | 2 | US-08-471-800-31 |
| 21 | 32 | 71.1 | 15 | 2 | US-08-488-161-20 |
| 22 | 32 | 71.1 | 15 | 2 | US-08-471-068-31 |
| 23 | 32 | 71.1 | 15 | 3 | US-09-273-685-20 |
| 24 | 32 | 71.1 | 15 | 5 | PCT-US95-11934-20 |
| 25 | 32 | 71.1 | 38 | 1 | US-08-176-500-22 |
| 26 | 32 | 71.1 | 38 | 1 | US-08-471-052A-22 |
| 27 | 32 | 71.1 | 38 | 1 | US-08-189-331-22 |

RESULT 1
US-08-272-255-8
; Sequence 8, Application US/08272255
; Patent No. 5824859
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; APPLICANT: Lin, Chentao
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/272,255
; FILING DATE: 08-JUL-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: UPN-1795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-272-255-8

ALIGNMENTS

Query Match 84.4%; Score 38; DB 2; Length 566;
Best Local Similarity 75.0%; Pred. No. 7, 7;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGMV 9

Db 88 LYDNVGLV 95


```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-145-6

Query Match 77.8%; Score 35; DB 1; Length 9;
Best Local Similarity 75.0%; Pred. No. 3e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

QY 2 LYENVGY 9
Db 1 LYQNVGY 8

RESULT 5
US-09-376-343-3
; Sequence 3, Application US/09376343
; Patent No. 6508592
; GENERAL INFORMATION:
; APPLICANT: Blum, Paul H.
; TITLE OF INVENTION: Hyperthermophilic Alpha-Glucosidase Gene and Its Use
; FILE REFERENCE: N1231-200
; CURRENT APPLICATION NUMBER: US/09/376,343
; CURRENT FILING DATE: 1999-08-18
; EARLIER APPLICATION NUMBER: 60/096,860
; EARLIER FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Sulfolobus solfataricus
US-09-376-343-3

Query Match 77.8%; Score 35; DB 4; Length 19;
Best Local Similarity 62.5%; Pred. No. 0.7;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVGY 9
Db 6 IYENLGVI 13

RESULT 6
US-08-480-190-38
; Sequence 38, Application US/08480190
; Patent No. 5827516
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; APPLICANT: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,379
; FILING DATE: August 11, 1992
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: 08/077,255
; APPLICATION NUMBER: 07/925,460
; FILING DATE: June 15, 1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-480-190-38

Query Match 77.8%; Score 35; DB 2; Length 20;
Best Local Similarity 75.0%; Pred. No. 0.74;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGY 9
Db 3 LYQNVGY 10

RESULT 7
US-08-488-379-38
; Sequence 38, Application US/08488379
; Patent No. 5880103
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,379
; FILING DATE: August 11, 1992
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: 08/077,255
; APPLICATION NUMBER: 07/925,460
; FILING DATE: June 15, 1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001

```

us-09-998-350-1.ra1

Tue Jul 20 16:13:37 2004

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 38:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-488-379-38

Query Match 77.8%; Score 35; DB 2;
 Best Local Similarity 75.0%; Pred. No. 0.74;
 Matches 6; Conservative 1; Mismatches 1;

QY 2 LYENVGMV 9
 DB 3 LYQNVGT 10

RESULT 8
 US-08-475-399A-38
 ; Sequence 38, Application US/08475399A
 ; Patent No. 6509033
 ; GENERAL INFORMATION:
 ; APPLICANT: Urban, Robert G.
 ; APPLICANT: Chic, Roman M.
 ; APPLICANT: Vignali, Dario A.A.
 ; APPLICANT: Hedley, Mary L.
 ; APPLICANT: Stein, Lawrence J.
 ; APPLICANT: Strominger, Jack L.
 ; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
 ; NUMBER OF SEQUENCES: 276
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson, P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: US
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/475,399A
 ; FILING DATE: 07-JUN-1995
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 08/077,255
 ; FILING DATE: 15-JUN-1993
 ; APPLICATION NUMBER: 07/925,460
 ; FILING DATE: 11-AUG-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fraser, Janis K.
 ; REGISTRATION NUMBER: 34,819
 ; REFERENCE/DOCKET NUMBER: 00246/168003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-507
 ; TELEFAX: 617/542-890
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 38:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 20 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-08-475-399A-38

Query Match 77.8%; Score 35; DB 4;
 Best Local Similarity 75.0%; Pred. No. 0.74;
 Matches 6; Conservative 1; Mismatches 1;

QY 2 LYENVGMV 9
 DB 3 LYQNVGT 10

RESULT 9
 PCT-US93-07545-38
 ; Sequence 38, Application PC/TUS9307545
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert G. Urban
 ; APPLICANT: Roman M. Chic
 ; APPLICANT: Dario A. A. Vignali
 ; APPLICANT: Mary L. Hedley
 ; APPLICANT: Lawrence J. Stern
 ; APPLICANT: Jack L. Strominger
 ; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
 ; NUMBER OF SEQUENCES: 273
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 50Z or 55SX
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)
 ; SOFTWARE: WordPerfect (Version 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/07545
 ; FILING DATE: 19930811
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 07/925,460
 ; FILING DATE: August 11, 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Clark, Paul T.
 ; REGISTRATION NUMBER: 30,162
 ; REFERENCE/DOCKET NUMBER: 00246/168001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 38:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 20
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; PCT-US93-07545-38

Query Match 77.8%; Score 35; DB 5; Length 20;
 Best Local Similarity 75.0%; Pred. No. 0.74;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
 DB 3 LYQNVGT 10

RESULT 10
 US-09-003-287-6
 ; Sequence 6, Application US/09003287
 ; Patent No. 6096947
 ; GENERAL INFORMATION:
 ; APPLICANT: Jayne, Susan
 ; APPLICANT: Barbour, Eric
 ; APPLICANT: Meyer, Terry
 ; TITLE OF INVENTION: METHODS FOR IMPROVING TRANSFORMATION EFFICIENCY
 ; FILE REFERENCE: MOPAT MOCAH
 ; CURRENT APPLICATION NUMBER: US/09/003,287
 ; CURRENT FILING DATE: 1998-01-06

```
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Myrothecium verrucaria
US-09-003-287-6

Query Match      77.8%; Score 35; DB 3; Length 244;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 LYENVGMY 9
      ||:|||||
Db      170 LYDNVGAY 177

RESULT 11
US-09-003-287-8
; Sequence 8, Application US/09003287
; Patent No. 6096947
; GENERAL INFORMATION:
; APPLICANT: Jayne, Susan
; APPLICANT: Barbour, Eric
; APPLICANT: Meyer, Terry
; TITLE OF INVENTION: METHODS FOR IMPROVING TRANSFORMATION EFFICIENCY
; FILE REFERENCE: mOPAT_mOCAH
; CURRENT APPLICATION NUMBER: US/09/003,287
; CURRENT FILING DATE: 1998-01-06
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Myrothecium verrucaria
US-09-003-287-8

Query Match      77.8%; Score 35; DB 3; Length 244;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 LYENVGMY 9
      ||:|||||
Db      170 LYDNVGAY 177

RESULT 12
US-09-518-988-2
; Sequence 2, Application US/09518988
; Patent No. 6268547
; GENERAL INFORMATION:
; APPLICANT: Weeks, James T.
; TITLE OF INVENTION: TRANSFORMATION OF WHEAT WITH THE
; TITLE OF INVENTION: CYANAMIDE HYDRATASE GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nancy J. Parsons
; STREET: 800 Buchanan St.
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/518,988
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/873,001

; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Myrothecium verrucaria
US-09-003-287-6

Query Match      77.8%; Score 35; DB 3; Length 244;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 LYENVGMY 9
      ||:|||||
Db      170 LYDNVGAY 177

RESULT 13
US-09-080-897-5
; Sequence 6, Application US/09080897
; Patent No. 5985574
; GENERAL INFORMATION:
; APPLICANT: King, Mary-Claire
; APPLICANT: Lynch, Eric D.
; APPLICANT: Lee, Ming
; APPLICANT: Morrow, Jan E.
; APPLICANT: Welsh, Piri L.
; APPLICANT: Leon, Pedro E.
; TITLE OF INVENTION: Modulators of Actin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,897
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UW97-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-080-897-6

Query Match      75.6%; Score 34; DB 2; Length 362;
Best Local Similarity 75.0%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      2 LYENVGMY 9
Db      248 LYENLGEY 255

RESULT 14
US-09-323-735-6
; Sequence 6, Application US/09323735
; Patent No. 6197932
; GENERAL INFORMATION:
; APPLICANT: King, Mary-Claire
; APPLICANT: Lynch, Eric D.
; APPLICANT: Lee, Ming
; APPLICANT: Morrow, Jan E.
; APPLICANT: Welcsh, Piri L.
; APPLICANT: Leon, Pedro E.
; TITLE OF INVENTION: Modulators of Actin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/323,735
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/080,897
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UM97-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-323-735-6

Query Match      75.6%; Score 34; DB 3; Length 362;
Best Local Similarity 75.0%; Pred.No. 31;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      2 LYENVGMY 9
Db      248 LYENLGEY 255

RESULT 15
US-09-252-991A-27339
; Sequence 27339, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
```

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27339
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27339

Query Match      73.3%; Score 33; DB 4; Length 310;
Best Local Similarity 71.4%; Pred.No. 42;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      3 YENVGMY 9
Db      205 YENAGLY 211

Search completed: July 20, 2004, 15:45:37
Job time : 19 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 20, 2004, 15:34:04 ; Search time 36 Seconds
(without alignments)
78.880 Million cell updates/sec

Title: US-09-998-350-1

Perfect score: 45

Sequence: 1 LYENVMGY 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 36 | 80.0 | 540 | 12 Q88452 | Q88452 sigma virus |
| 2 | 36 | 80.0 | 848 | 5 Q813V8 | Q813V8 plasmodium |
| 3 | 36 | 80.0 | 914 | 12 Q91F30 | Q91F30 bovine aden |
| 4 | 36 | 80.0 | 1525 | 3 Q96W88 | Q96W88 kluyveromyc |
| 5 | 35 | 77.8 | 339 | 12 Q91FF9 | Q91FF9 influenza a |
| 6 | 35 | 77.8 | 339 | 12 Q91FF4 | Q91FF4 influenza a |
| 7 | 35 | 77.8 | 339 | 12 Q91FF2 | Q91FF2 influenza a |
| 8 | 35 | 77.8 | 339 | 12 Q91FF6 | Q91FF6 influenza a |
| 9 | 35 | 77.8 | 339 | 12 Q91FF0 | Q91FF0 influenza a |
| 10 | 35 | 77.8 | 339 | 12 Q91FF7 | Q91FF7 influenza a |
| 11 | 35 | 77.8 | 339 | 12 Q91FF8 | Q91FF8 influenza a |
| 12 | 35 | 77.8 | 339 | 12 Q91FF5 | Q91FF5 influenza a |
| 13 | 35 | 77.8 | 339 | 12 Q91FF0 | Q91FF0 influenza a |
| 14 | 35 | 77.8 | 339 | 12 Q91FF3 | Q91FF3 influenza a |
| 15 | 35 | 77.8 | 339 | 12 Q91FF1 | Q91FF1 influenza a |
| 16 | 35 | 77.8 | 339 | 12 Q91FF2 | Q91FF2 influenza a |

| | | | | | |
|----|----|------|-----|-----------|--------------------|
| 17 | 35 | 77.8 | 339 | 12 Q91FG1 | Q91FG1 influenza a |
| 18 | 35 | 77.8 | 359 | 12 Q997B2 | Q997B2 influenza a |
| 19 | 35 | 77.8 | 359 | 12 Q997B3 | Q997B3 influenza a |
| 20 | 35 | 77.8 | 359 | 12 Q997B4 | Q997B4 influenza a |
| 21 | 35 | 77.8 | 359 | 12 Q997B1 | Q997B1 influenza a |
| 22 | 35 | 77.8 | 373 | 12 Q9WQX2 | Q9WQX2 influenza a |
| 23 | 35 | 77.8 | 376 | 12 Q9WQX1 | Q9WQX1 influenza a |
| 24 | 35 | 77.8 | 376 | 12 Q9WQW1 | Q9WQW1 influenza a |
| 25 | 35 | 77.8 | 376 | 12 Q9WQW4 | Q9WQW4 influenza a |
| 26 | 35 | 77.8 | 378 | 12 Q9WQX0 | Q9WQX0 influenza a |
| 27 | 35 | 77.8 | 378 | 12 Q9WQW8 | Q9WQW8 influenza a |
| 28 | 35 | 77.8 | 378 | 12 Q9WQW6 | Q9WQW6 influenza a |
| 29 | 35 | 77.8 | 378 | 12 Q9WQW2 | Q9WQW2 influenza a |
| 30 | 35 | 77.8 | 379 | 12 Q9WQX3 | Q9WQX3 influenza a |
| 31 | 35 | 77.8 | 380 | 12 Q9WQV9 | Q9WQV9 influenza a |
| 32 | 35 | 77.8 | 381 | 12 Q9WQW7 | Q9WQW7 influenza a |
| 33 | 35 | 77.8 | 381 | 12 Q9WQW5 | Q9WQW5 influenza a |
| 34 | 35 | 77.8 | 381 | 12 Q9WQW3 | Q9WQW3 influenza a |
| 35 | 35 | 77.8 | 381 | 12 Q9WQW0 | Q9WQW0 influenza a |
| 36 | 35 | 77.8 | 448 | 16 Q51669 | Q51669 borrelia bu |
| 37 | 35 | 77.8 | 560 | 12 Q9WQW9 | Q9WQW9 influenza a |
| 38 | 35 | 77.8 | 562 | 12 Q67032 | Q67032 influenza a |
| 39 | 35 | 77.8 | 562 | 12 Q67085 | Q67085 influenza a |
| 40 | 35 | 77.8 | 562 | 12 Q67208 | Q67208 influenza a |
| 41 | 35 | 77.8 | 562 | 12 Q67120 | Q67120 influenza a |
| 42 | 35 | 77.8 | 562 | 12 Q67011 | Q67011 influenza a |
| 43 | 35 | 77.8 | 562 | 12 Q67284 | Q67284 influenza a |
| 44 | 35 | 77.8 | 562 | 12 Q67165 | Q67165 influenza a |
| 45 | 35 | 77.8 | 562 | 12 Q67143 | Q67143 influenza a |

ALIGNMENTS

RESULT 1

Q88452 ID Q88452 PRELIMINARY; PRT; 540 AA.
AC Q88452;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE glycoprotein.
GN G.
OS Sigma virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; unclassified Rhabdoviridae.
OX NCBI_TaxID=11301;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=234HRC;
RX MEDLINE=96074506; PubMed=7491755;
RA Landes-Devauchelle C., Bras F., Dezelee S., Teninges D.,
RT "Gene 2 of the sigma rhabdovirus genome encodes the P protein, and
RT gene 3 encodes a protein related to the reverse transcriptase of
RT retroelements.";
RL Virology 213:300-312(1995).
DR EMBL; X91062; CAA62517.1; -;
DR InterPro: IPR001903; Rhabd glycop.
DR Pfam: PF00574; Rhabdo glycop; 1
SQ SEQUENCE 540 AA; 60771 MW; 7A0B553D1EASE98A CRC64;

Query Match Best Local Similarity 80.0%; Score 36; DB 12; Length 540;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LYENVMGY 9

Db 365 LYQSVGMY 372

RESULT 2

Q813V8 ID Q813V8 PRELIMINARY; PRT; 848 AA.


```

OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=220951;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Davis/1/57;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270719; AAF82103.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Glycoprotein; Hemagglutinin.
KW NON TER 339
FT SEQUENCE 339 AA; 37810 MW; 7D35925ED753B08 CRC64;
SQ

Query Match 77.8%; Score 35; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
DB 204 LYQNVGT 211
||:|||||
|:|:|:|

RESULT 6
Q9IFF4 PRELIMINARY; PRT; 339 AA.
ID Q9IFF4
AC Q9IFF4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Malaya/16/58 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=220954;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Malaya/16/58;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270724; AAF82108.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Glycoprotein; Hemagglutinin.
KW NON TER 339
FT SEQUENCE 339 AA; 37893 MW; D59A261E1EB9B621 CRC64;
SQ

Influenza A viruses; Influenzavirus A.
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
DB 204 LYQNVGT 211
||:|||||
|:|:|:|

RESULT 7
Q9IFF2 PRELIMINARY; PRT; 339 AA.
ID Q9IFF2
AC Q9IFF2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Victoria/15681/59 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=220956;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Victoria/15681/59;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270726; AAF82110.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR InterPro; IPR001364; Hemagglutn.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
DR Envelope protein; Glycoprotein; Hemagglutinin.
KW NON TER 339
FT SEQUENCE 339 AA; 37964 MW; 97239D60CD1FFD08 CRC64;
SQ

Query Match 77.8%; Score 35; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
DB 204 LYQNVGT 211
||:|||||
|:|:|:|

RESULT 8
Q9IFF6 PRELIMINARY; PRT; 339 AA.
ID Q9IFF6
AC Q9IFF6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/RI/5+/57 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135328;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/RI/5+/57;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
```

```

RT avian influenza virus hemagglutinins after their introduction into
RL mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270722; AAF82106.1; -.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR InterPro; IPR001364; Hemagglutn.
DR PRINTS; PR00329; Hemagglutn12.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON_TER 339 339
SQ SEQUENCE 339 AA; 37853 MW; 7C70576EBB5B2EC0 CRC64;

Query Match 77.8%; Score 35; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
DB 204 LYQNVGT 211

RESULT 9
Q9IF60
ID Q9IF60 PRELIMINARY; PRT; 339 AA.
AC Q9IF60
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Chile/6/57 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Chile/6/57;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270728; AAF82112.1; -.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR InterPro; IPR001364; Hemagglutn.
DR PRINTS; PR00329; Hemagglutn12.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON_TER 339 339
SQ SEQUENCE 339 AA; 37810 MW; 7D35925ED7538B08 CRC64;

Query Match 77.8%; Score 35; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
DB 204 LYQNVGT 211

```

```

RESULT 10
Q9IFF7
ID Q9IFF7 PRELIMINARY; PRT; 339 AA.
AC Q9IFF7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (strain A/Ann Arbor/6/60).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135322;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Ann Arbor/6/60;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270721; AAF82105.1; -.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR InterPro; IPR008980; Capsid hemag.
DR InterPro; IPR001364; Hemagglutn.
DR PRINTS; PR00329; Hemagglutn12.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON_TER 339 339
SQ SEQUENCE 339 AA; 37896 MW; FECE7718D2628F0E CRC64;

Query Match 77.8%; Score 35; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LYENVGMV 9
DB 204 LYQNVGT 211

RESULT 11
Q9IFF8
ID Q9IFF8 PRELIMINARY; PRT; 339 AA.
AC Q9IFF8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Hemagglutinin (Fragment).
OS Influenza A virus (A/Albany/7/57 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135321;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Albany/7/57;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.

```



```

DR EMBL; AF270720; AAF82104.1; -
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid_hemag.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37825 MW; 0D3E767F9241AA30 CRC64;

Query Match 77.8%; Score 35; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LYENVGMV 9
Db 204 LYQNVGT 211

RESULT 12
Q9IFFS PRELIMINARY; PRT; 339 AA.
AC Q9IFFS
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment)
OS Influenza A virus (A/Albany/6/58 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=220953;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Albany/6/58;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270718; AAF82102.1; -
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid_hemag.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37798 MW; FE7698C4DC1D15E6 CRC64;

Query Match 77.8%; Score 35; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LYENVGMV 9
Db 204 LYQNVGT 211

RESULT 14
Q9IFFS PRELIMINARY; PRT; 339 AA.
AC Q9IFFS
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hemagglutinin (Fragment)
OS Influenza A virus (A/Sao Paulo/3/59 (H2N2)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135330;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Sao Paulo/3/59;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawaoka Y.;
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270725; AAF82109.1; -
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR008980; Capsid_hemag.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; Hemagglutinin; 1.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON TER 339
SQ SEQUENCE 339 AA; 37753 MW; 2ADC4BA8C590ADCE CRC64;

Query Match 77.8%; Score 35; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LYENVGMV 9
Db 204 LYQNVGT 211

RESULT 13
Q9IFGO PRELIMINARY; PRT; 339 AA.
AC Q9IFGO
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

FT NON_TER 339 339
SQ SEQUENCE 339 AA; 37895 MW; 97D69D60CD5AFD08 CRC64;
Query Match 77.8%; Score 35; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LYENVGMV 9
||:||||
Db 204 LYQNVGT 211

RESULT 15
Q9IFF1 PRELIMINARY; PRT; 339 AA.
AC Q9IFF1
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hemagglutinin (Fragment)
OS Influenza A virus (A/Oslo/2/59 (H2N2))
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=135327;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/Oslo/2/59;
RA Matrosovich M., Tuzikov A., Bovin N., Gambaryan A., Klimov A.,
RA Castrucci M.R., Donatelli I., Kawacka Y.,
RT "Early alterations of the receptor-binding properties of H1, H2 and H3
RT avian influenza virus hemagglutinins after their introduction into
RT mammals.";
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY). CHAINS
CC -!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
DR EMBL; AF270727; AAF82111.1;
DR GO; GO:0019031; C: viral envelope; IEA.
DR InterPro; IPR008980; Capsid_hemag.
DR Pfam; PF00509; Hemagglutinin; 1.
DR PRINTS; PR00329; HEMAGGLUTN12.
DR ProDom; PD000225; Hemagglutn; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON_TER 339
SQ SEQUENCE 339 AA; 37991 MW; F6BC8A0403FD40CC CRC64;

Query Match 77.8%; Score 35; DB 12; Length 339;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LYENVGMV 9
||:||||
Db 204 LYQNVGT 211

Search completed: July 20, 2004, 15:44:41
Job time : 38 secs